Cheater-detection, Deceit-detection or Task Understanding?

Studies on the Wason Selection Task.

A thesis

submitted in partial fulfilment

of the requirements for the Degree

of

Masters of Arts in Psychology

in the

University of Canterbury

by

Jonathan M. Grady

University of Canterbury

1998

Acknowledgements:

Many, many thanks to all the people that helped and supported me through this thesis. Special thanks go to:

Paul Russell, my supervisor;
Jack Copeland, my philosophical mentor;
Russil Durrant and **Dave Hardman** for their theoretical input;
Kathy Keenan for the fabulous pictures used in Experiment 2;
Conal Grady for the data collection for Experiment 2;
Simon Horsburgh and **Tom Keenan** for all the coffee;
Gill Rhodes and **Chris Burt** for the employment;
Mum & Dad for their tolerance;
Simon Kemp for the encouragement;
Tess, **Meg**, **Kaz**, **Paul B.**, **Colette**, & **Peter** for the sanity and insanity;
The Participants (all 300 of you);
& **Salvador Dali.**

Many Thanks, Jonny.

Contents:

Abstract

A number of theories of dealing with performance on the Wason Selection Task were examined. Experiments 1a and 1b tested Cosmides' (1989) Social Exchange theory, which appeals to "cheater-detection", against a more general "deceit-detection" hypothesis. Three social contract tasks, which cued the participants to the perspective of an individual who was looking for instances of cheating on a social contract, were tested against three non-social-contract tasks, which cued the participants to the perspective of an individual who was looking for instances of lying. The initial results appeared to support Social Exchange theory and reject the "deceit-detection" hypothesis. However, the results were also explicable in terms of Liberman & Klar's (1996) Theory of Task Understanding.

Experiment 2 tested the Theory of Task Understanding against Social Exchange theory. Three new social contract tasks were tested against three thematic-abstract tasks, which were designed in accordance with prescriptions that Liberman & Klar's theory suggests should lead to good task performance. A small, significant difference was found between the social contract tasks (50% correct) and the thematic-abstract tasks (40%), which is consistent with the prediction of Social Exchange theory. However, a more detailed analysis revealed that the within-condition range for both conditions was three times larger than the between-condition difference. This is problematic for Social Exchange theory, but consistent with the Theory of Task Understanding. Experiment 2 also tested whether the inclusions of pictures or diagrams, that represent the problem, increase task performance. No significant difference existed between tasks that included pictures and those that did not.

1 General Introduction

1.0 Welcome

Thought and, in particular, the act of hypothetical reasoning are among the most powerful tools in the human cognitive repertoire. They provide us with the ability to deal with abstract as well as concrete problems, and enable us to predict the outcomes of many events before they occur. Without these abilities, many problems would be insurmountable, and our culture and technology could never have developed.

The aim of this thesis is to investigate a number of theories, which attempt to explain performance on the *Wason Selection Task* – a powerful tool for research into *deductive reasoning*. To begin with, though, a number of definitions are required to map out the domain this thesis concerns itself with.

Reasoning has been grouped into two separate domains: Deduction and Induction. Evans (1982) suggests that deduction can be thought of as leading thinking from the *general* to the *particular* as in:

(1) All grass is green.

(2) This lawn is made of grass.

(C) Therefore, this lawn is green.

Induction, on the other hand, leads from the *particular* to the *general* as in:

(1) All the grass I have ever seen is green.

(C) All grass is green.

Deduction is concerned with drawing conclusions from a given set of facts, while induction involves compiling together various instances of an event and inferring an integrative conclusion. It is with the field of *deduction* that this thesis concerns itself.

1.1 Deduction

Deduction is the process of drawing a logically valid[1] conclusion from a set of given statements or 'premises'. One takes two or more statements of fact (which are assumed to be true) that are in linked by content, and attempts to find a derivable outcome for the case when all the statements occur. For instance, given the following two premises:

(1) When Billy has a headache, he is grumpy.

(2) Billy has a headache.

---

[1] A conclusion is logically **valid** if a series of assumptions which are true **cannot** lead to a conclusion which is false (Evans, 1982).

We can deductively conclude the following valid conclusion:

(C) Billy is grumpy.

In this case we have been given an initial statement (1) which describes a relationship between two propositions, and another statement (2) which tells us that one of the propositions has occurred. The conclusion (C) is the most parsimonious statement that can occur when both premises are taken into account. It should be noted that the statements do not have to be true in the real world for the conclusion to be valid. Likewise, not every conclusion that is true in the real world is logically valid either. Take this example (from Oakhill & Johnson-Laird, 1985):

(1) All of the Frenchmen are wine drinkers.

(2) Some of the wine drinkers are gourmets.

(C) Some of the Frenchmen are gourmets.

The conclusion (C) here is invalid[1] even though we know it is true in the real world.

Comparison of the two previous examples also brings up another point. Deductive logic problems can take very different forms from one another. Note that the first example (Billy's headache) is an "If X Then Y..." type argument while the second one (the French wine drinkers) is an "All X are Y..." type argument. The first form is an example of *propositional logic* (so called because it involves analysis of a string of propositions), while the second form is an example of *quantitative logic* (so called because the problem involves comparing quantitative terms such as "All", "Some" or "None"). A third variety of logic problem is known as *relational logic*, so called because it involves analysis of relationships between the properties of the items concerned (Evans et al., 1993). An example of this would be:

(1) Sally is taller than Bob.

(2) Bob is taller Lucy.

(C) Sally is taller than Lucy.

Here the relationship in question is the relative heights of the people concerned, but we could just as easily be examining other properties such as physical position ("Sally is to the right of Bob, Lucy is to the left of Bob...") or intelligence ("Lucy is smarter than Sally...").

Of the three types of deduction, this thesis is only concerned with *propositional logic.*

---

[1] To see why this is invalid, try substituting "Italians" for "gourmets".

1.2 Propositional Logic

Propositional logic involves examining the causal structure of two or more linked propositions and, given information about the *truth-value*[1] of one of the propositions, finding a valid conclusion. Consider the following propositions:

(i) You stand on a spider.

(ii) It is raining.

These propositions can each have two truth-values – true or false. In the case of (i) if you are standing on a spider then the statement is *true*. If you are not standing on a spider the statement is *false*. Likewise, with (ii) if it is raining the statement is *true*, if it is not raining the statement is *false*. The two propositions may be causally linked by the following structure:

(1) If you stand on a spider, it will rain.

In order to create a full argument, additional information about the truth-value of one of the propositions is required:

(2) You have stood on a spider.

Now, because the premise (1) states that a true instance of proposition (i) will necessarily lead to a true instance of proposition (ii), we can draw the valid conclusion:

(C) It will rain.

Such a conclusion is known as *Modus Ponens*, and can be notated in this abstract way:

If P ⇒ Q	(*If* P *then* Q)[2]	
P	(*Given* P)	
∴ **Q**	(*Therefore*, Q)[3]	(Where P and Q are separate propositions).

It should be noted here that the "If...then..." statement is *unidirectional* from "P" to "Q". This means that while a true instance of "P" will necessarily lead to a true instance of "Q", a true instance of "Q" does not necessarily lead to a true instance of "P". Take the example above again. While it is true (hypothetically) that standing on a spider will cause it to rain, the fact that it may be raining does not mean that you have stood on a spider. There may be many alternative causes for the consequent "Q" without the need for the given antecedent "P". The misinterpretation of the "If" in the "If...then..."

---

[1] By "truth-value" we mean "did the proposition occur or not?" The proposition is *true* if it did occur, or *false* if it did not occur. It says nothing about the truthfulness of the proposition in the real world.

[2] A semantic equivalent to this is "P only if Q". Although logically equivalent, researchers have found that when given as problems, people tend to respond slightly differently to these two forms (Evans et al., 1993).

[3] A three line conditional problem such as this, consisting of two premises and a conclusion, is known as a *conditional syllogism*. Larger problems consisting of more premises and more conclusions have been examined in the literature (see Rip's 1989 "Knights and Knaves" problems for example) but this thesis is confined to conditional syllogisms only.

statement as meaning "If-and-only-if" (known as the *biconditional* "if") often leads people to accept invalid conclusions to problems of this form (Anderson, 1990; Evans et al., 1993).

As well as Modus Ponens, there is another valid conclusion that can be derived from a conditional syllogism. Consider at the above example again:

(1) If you stand on a spider, it will rain.

Now, instead of assigning a *true* instance of the *first proposition* (the antecedent), contemplate what happens when a *false* (or negative) instance of the *second proposition* (the consequent) is inserted into the second premise.

(2) It is *not* raining.

From these premises we can validly conclude:

(C) You have not stood on a spider.

The proof of this conclusion is not immediately obvious to many people. From (1) we know that if the antecedent is true (you stand on a spider) then the consequence will be true (it will rain). But from (2) we know that the consequence is false (it is not raining), so the antecedent cannot be true. This leaves us with the only alternative, which is that the antecedent is false (you have not stood on a spider). This conclusion is known as *Modus Tollens*. It can be expressed abstractly as:

If P $\Rightarrow$ **Q** (*If* P *then* Q)
~Q (*Given* not Q)
$\therefore$ **~P** (*Therefore*, not P)

Modus Ponens and Modus Tollens are the only valid conclusions that can be derived from a conditional syllogism. There are two other arguments that people often accept as valid, even though they are not. These are:

If P $\Rightarrow$ **Q**
~P
$\therefore$ **~Q**

known as *Denial of the Antecedent*, and

If P $\Rightarrow$ **Q**
Q
$\therefore$ **P**

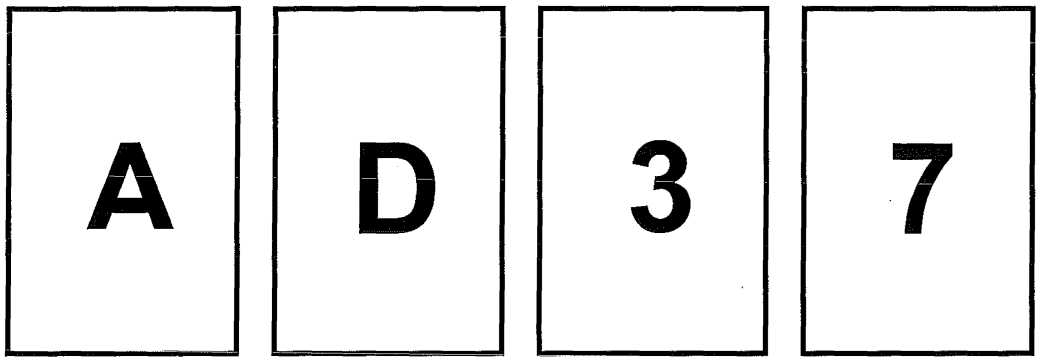
known as *Affirmation of the Consequent* (Anderson, 1990; Evans et al., 1993). The fact that people often accept these fallacies as valid arguments is probably at least partly due to the misinterpretation of the "If" as biconditional (Taplin & Staudenmayer, 1973).

Now that the philosophical basics of propositional reasoning have been outlined, an examination of psychological research concerning conditional syllogisms can be presented.

1.3 The Wason Selection Task

Way and by far, the most extensively researched problem in psychological reasoning is Peter Wason's (1966) four card selection task, or *the Wason Selection Task* (Evans et al., 1993). The Wason Selection Task is a tool for presenting participants with conditional syllogisms framed within a contextual problem. The example below, taken from Wason & Johnson-Laird (1972), shows a standard abstract version of the task:

> *Subjects are shown a set of cards, each of which is seen to have a capital letter on one side and a single figure number on the other side. The experimenter then hides the cards and selects four which are placed on the table* [in random order]. *The subject can then see the four facing sides as follows:*



> *The subject is then told that "the following rule applies to these four cards and may be true or false"*
>
> ***If there is an A on one side of the card,***
> ***then there is a 3 on the other side of the card.***
>
> *The subject is then asked to decide which of the four cards would need to be turned over in order to decide whether the rule is true or false.*

In this task the "rule" can be viewed as the first premise in a conditional syllogism – the causal structure linking two propositions. Each card can be viewed as the second premise in the syllogism – the given truth-value one of the propositions. Hence, the task requires the participant to evaluate four separate conditional syllogisms, one for each card. The "A" card in this example is the equivalent of **P** (a true instance of the first proposition), so turning it over (to make sure there is a "3" on the other side) is the equivalent of Modus Ponens. The "D" card is the equivalent of **~P** (a false instance of

the first proposition), so turning it over is the equivalent of Denial of the Antecedent. The "3" card is equivalent to **Q** (a true instance of the second proposition), so turning it over is the equivalent of Affirmation of the Consequent. Finally, the "7" card is the equivalent of **~Q** (a false instance of the second proposition), so turning it over (to make sure there is not an "A" on the other side) is the equivalent of Modus Tollens. Since the only valid conclusions to a conditional syllogism are Modus Ponens and Modus Tollens, we can see that the correct answer to the Wason Selection Task is **P & ~Q** – in this case turning over the "A" and "7" cards. If the "A" card is turned over and a number *other* than "3" is found on the back, then the rule has been falsified. Likewise, if the "7" card is turned over and found an "A" is found on the back, this also falsifies the rule. On the other hand, turning over the "D" or "3" cards in this case does not test the "rule" at all. The rule does not say anything about what we should expect to find on the back of a "D" (or any other letter other than "A"), nor does it prohibit any letter other than "A" being on the back of a "3" card.

In this abstract version of the selection task people typically indicate that the "A" alone (**P only**) should be turned over (about 30-35%) or that "A" and "3" (**P & Q**) should be turned over (about 45%). Both of these answers are wrong. In fact, only about 4-10% of participants correctly elect to turn over the "A" and "7" cards (**P & ~Q**) (Wason & Johnson-Laird, 1972). These results highlight two dominant findings in the literature concerning abstract versions of the task: firstly, people exhibit a bias towards selecting the unnecessary **Q** card, and secondly, they show a bias toward neglecting the necessary **~Q** card (Evans, 1982; Evans et al., 1993; Rips, 1994).

These biases seem quite resistant to factors such as intelligence, education and training. Valentine (1975, in Evans et al., 1993) examined the relationship between IQ and task performance and found a low but significant positive correlation. Jackson and Griggs (1988) examined level of education and task performance for university students from a number of different areas of expertise. They found no significant effects except that participants who were mathematics majors tended to perform better than other students. And, rather ironically, Cheng, Holyoak, Nisbett and Oliver (1986) tested participants who had received a university course of formal training in logic and found that they performed the selection task no better than a control group with no formal training.

Originally, these findings were thought to be attributable to *confirmation bias*[1] – the tendency for participants to attempt to prove the rule correct as opposed to attempting to falsify the rule. This notion was supported by the fact that, even when subjects were found to understand the falsifying potential of the **P & ~Q** combination, many refused to change their answers from the **P & Q** response (Evans et al., 1993). But another tendency, *matching bias*[2], has been shown to explain the results as well, if not better. Simply stated, matching bias is *the tendency to pick cards corresponding to the stated propositions in the initial rule* (Evans & Lynch, 1973; Evans, 1995). Matching bias is best demonstrated when negatives are inserted into the rule. Compare the following two examples.

As described earlier, it is well established that given:

"If A then B" (**P⇒Q**)

more people will choose the "B" card (**Q**) than the "not B" (**~Q**) card. But it has also been found that when given:

"If A then not B" (**P⇒Q**)

more people will still choose the "B" card (**~Q**) than the "not B" card (**Q**), even though the consequent (**Q**) has been changed. If people were adhering to confirmation bias, then they would be expected to select the **Q** card regardless of whether it was negated.

In the case above, the change *improves* the level of correct responding to the task. However, matching bias can also occur on the antecedent. If the antecedent is negated ("If not A then B"), then more people still tend choose the "A" card, even though it is now incorrect.

1.4 The Thematic Content Effect

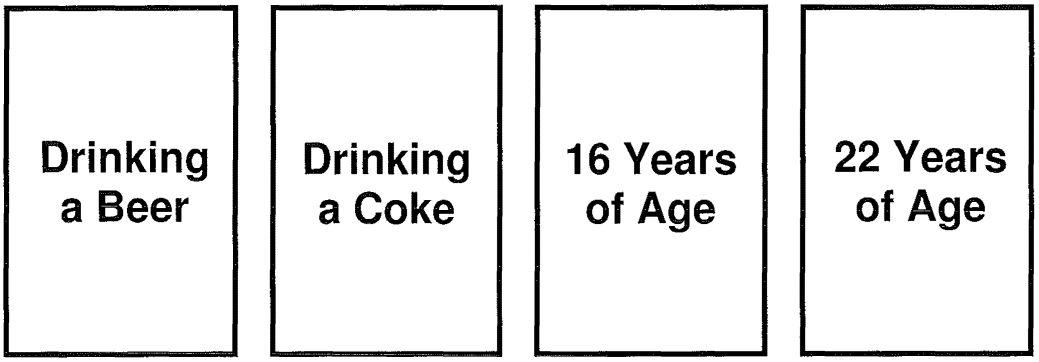
The peculiarities of the Wason Selection Task do not end with the abstract versions of the task. It has been found that the introduction of various types of thematic content differentially facilitates correct responding to the problem. The thematic content is introduced as a short scenario which cues the participant to the perspective of someone checking to see if a given rule is being followed or falsified. The selection task is imbedded within the scenario. The facilitation that can be achieved with this sort of thematic content is well exhibited in Griggs & Cox's (1982) famous "Drinking Age rule":

---

[1] Although, this only accounted for the **P & Q** result, not the **P only** result. It does, however, go some way to explaining the low rates of **~Q** responses.

[2] The explain of why matching bias has overshadowed confirmation bias as the most likely contributor to fallacious reasoning on the abstract version of the Wason Selection Task is a too long winded to be shown here in its entirety. Any reader interested in learning more about this, though, should refer to Evans & Lynch (1973).

On this task imagine you are a police officer on duty. It is your job to ensure that people conform with certain rules. The cards in front of you have information about four people sitting at a table. On one side of a card is a person's age and on the other side is what a person is drinking. Here is a rule: "IF A PERSON IS DRINKING BEER, THEN THE PERSON MUST BE OVER 19 YEARS OF AGE". Select the card, or cards that you definitely need to turn over to determine whether or not people are violating the rule.



Griggs and Cox found that this form of the selection task elicited a massive 74% correct responses for "Drinking Beer" and "16 Years of Age" (**P & ~Q**). Compare this to the 4-10% found on abstract versions of the task. This is typical for thematic versions: generally the level of correct responding is around 65-80% (for a general review see Evans et al., 1993), although success rates as high as 96% (Gigerenzer & Hug, 1992) have been found in one-off samples. Much lower success rates have also been recorded: a few studies have found success rates at around 40-50%, and some show no facilitation at all (Manktelow & Evans, 1979[1], in Evans et al., 1993).

As well as facilitation of correct responding, the effect of thematic content also produces some interesting side effects. Firstly, the matching bias found in the purely abstract versions of the task seems to disappear in thematic content versions (except for Manktelow & Evan's example above). Secondly, when abstract versions are presented to participants along with thematic versions, there seems to be a slight level of facilitation, generally to about a 20% success rate. This indicates that some participants may use a form of reasoning-by-analogy from the thematic to the abstract if they can see the structural similarities of the two forms of the task. These two points, coupled with the thematic facilitation effect suggest that people may reason quite differently between abstract and thematic versions of the task. Some researches suggest that participants may use entirely different reasoning processes in dealing with

[1] Although it contained some thematic content, Manktelow & Evans' example ("*If I eat haddock, then I drink gin*") contains such arbitrary propositions that the participants probably treated it as a purely abstract task (Evans et al., 1993).

these problems, while some others suggest that people may not be reasoning at all per se (Evans et al., 1993).

Now that the main findings of the Wason selection task have been summarised, a brief review of some of the most important attempts to explain the results is presented. Although many of these explanations were specifically formulated with the selection task in mind, they are, of course, attempts to explain conditional reasoning in general.

1.5 Theories of Reasoning on the Wason Selection Task

The many theories that try to explain reasoning on the selection task can be grouped into two distinct categories – *Global theories* and *Domain-Specific theories*. Global theories suggest that all conditional reasoning, including mistakes and biases, can be explained using the same principles and mechanisms. Domain-Specific theories take the fact that people give different response patterns for different versions of the selection task as evidence that they use different principles and mechanisms under different domains. This section begins by examining a few of the global theories, then moves on to a couple of the domain-specific ones.

The first of the global theories is known as *Formal Rules of Inference* (sometimes known as "*Mental Proofs Theory*"). In fact, there are a number of different formulations of this theory (for instance Braine & O'Brien, 1991; Taplin & Staudenmayer, 1973; Rips, 1983) but all contain the following characteristics. Firstly, they suggests that people formulate, through direct experience of causes and outcomes, a set of implicit mental axiomatic rules. These implicit rules are short, proposition-like units designed for each of the logical connectives (such as "and", "or", "if", "then" and "not"). Secondly, when faced with a deductive problem, people attempt to uncover the logical form of the premises, identifying the underlying propositions and connectives. They remove the propositional content, and work with the underlying structure of the connectives. Thirdly, reasoners utilise the implicit mental rules to construct a mental derivation or proof of conclusion. Finally, they substitute the propositional content back into the conclusion. In the case of Modus Ponens, people already have an implicit rule that encapsulates the problem:

If A then B

So, given the first proposition we can directly reason to the conclusion:

A

$\therefore$ **B**

Thus, the derivation requires only one step (or three "lines").

The Formal Rules hypothesis suggests that many of the errors found on the abstract selection task arise because some derivations take a greater number of steps than others. This correctly predicts the problem people have with Modus Tollens, because the derivation requires 5 steps:

If A then B

~B (given)

A (generated by supposition or hypothesis)

∴**B** (from first line and third line)

∴**B and ~B** (from second and fourth lines)

reductio ad absurdum (the fifth line is rejected and, therefore, so is the third line)

∴**~A** (the negation of the third line. Hence, Modus Tollens).

The Formal Rules theory also has had some success with other area of abstract propositional reasoning and even within relational and quantified reasoning (Rips, 1994). But it completely fails to predict the thematic content effect on the selection task. According to Formal Rules theory, the reasoning should be oblivious to the content of the problem (Evans et al., 1993; Johnson-Laird, 1993; Johnson-Laird & Byrne, 1991).

The second global theory examined is *Mental Model Theory* (Johnson-Laird, 1993; Johnson-Laird & Byrne, 1991). This theory suggests that people construct detailed mental models of the problems they encounter and then attempt to evaluate which models are allowable given the truth-values of the propositions. The process of reasoning has three distinct stages:

(1) *Comprehension*: the reasoner attempts to make sense of the problem by constructing a mental model that describes the initial state of affairs. Each entity in the problem is represented by a "token" in the mental model. The properties of each entity are represented by the properties of each corresponding token. Relations between entities are represented by relations between their tokens. For example, take the relational premise "The square is to the right of the triangle". This could be expressed as a mental model like this:

Here the square is represented by a square token, the triangle is represented by a triangular token, and the translational relationship between the two is represented by the position of the tokens. The model may be expressed as a mental image in many cases, but Johnson-Laird and Byrne claim that this is not necessary. They also point out that many properties of tokens cannot be visualised (like negation, for example) so mental imagery may not be appropriate in some cases. Internal verbal representations

may also be used. What is important is not the manner of the subjective experience, but that the structure of the model represents the premise.

(2) *Description*: the reasoner searches the initial mental model in an attempt to formulate a putative conclusion to the problem. This conclusion should assert something that is not explicitly stated in the original premises and be parsimonious. If the person cannot determine such a conclusion, they will respond that no conclusion exists.

(3) *Validation*: the reasoner then attempts to construct alternative models of the premises in which the putative conclusion is false[1]. If no such model is found, then the reasoner accepts the putative conclusion as valid. If there is a falsifying model, then the reasoner should go back to (2) again and repeat the process. The set of possible models should be exhaustively searched.

Johnson-Laird and Byrne go on to claim that most of the errors on the Wason Selection Task (and, indeed, all reasoning) can be attributable to two main factors. Firstly, the larger the number of models needed to completely flesh out the implicit model, the harder the problem will be. More models place a greater load on working memory (Baddeley, 1986). Hence, with a greater number of models needed, the chances of any particular model being displaced from working memory or forgotten is higher. Secondly, as the set of possible models gets larger, the lower the chance is that people will construct and evaluate a valid model. Again, this is because of the limited resources of working memory.

Mental Model theory correctly predicts many of the results found not only in propositional reasoning, but relational and quantified[2] reasoning as well. In fact Evans, Newstead & Byrne's (1993) review of the psychological literature concerning deductive reasoning concluded that Mental Models was probably the most powerful and versatile theory on offer at the time.

With regards to the selection task, Johnson-Laird and Byrne account for the failure to select the **~Q** card in much the same way as Formal Rules theory. While the Modus Ponens solution requires only the initial model to be used as a putative conclusion, the Modus Tollens solutions requires the fleshing out of another three models as well (Johnson-Laird and Byrne, 1991, pg.48).

---

[1] Johnson-Laird Byrne call this "*fleshing out the implicit model*". The "implicit model" is really the set of models that are allowed by the premises but are not explicitly stated in the premises.

[2] The Mental Models theory was, in fact, first devised to account for quantified reasoning.

The inattention to fleshing out the implicit model is also the supposed reason for people falsely selecting the **Q** card - they do not construct and examine the correct falsifying model. Johnson-Laird and Byrne's attempt to account for the thematic content effect is a little vague, though. Again, they appeal to the fleshing out of the implicit model, saying that this should be easier under certain thematic conditions. They fail to suggest why the fleshing out process should be easier, although they appear to be alluding to the utilisation of general heuristics such as *availability* and *relevance* (which are examined next). While appearing somewhat unclear, this does make some sense: if we are familiar with certain material, we tend to be able to fit more of it into working memory (Anderson, 1990; Baddeley, 1986). Hence, the more familiar we are with the material, the more models we may be able to evaluate.

The third global theory reviewed is Evans' *Heuristic-analytic theory* (Evans, 1984; Evans, 1995). Rather than being a full theory in itself, heuristic-analytic theory is more a philosophy which incorporates a number of heuristic devices. The theory suggests that reasoning occurs in two distinct stages. In the first stage we select what seem to be the most obvious answers to our problem by utilising various unconscious, non-verbal heuristic processes. The second stage involves application of (unspecified)[1] analytic cognitive reasoning processes that evaluate choices made in the heuristic stage. More often than not, this analytic stage serves only to rationalise the heuristically generated choices in the context of the problem's instructions (Evans, 1995). Therefore, Evans suggest that often we do not really reason at all, but rather utilise "rules of thumb" to answer the problem. A number of heuristics have been mooted in the literature, but only the *availability* and *the relevance heuristics* are examine here (as these are the most cited and most researched).

The availability heuristic was originally born out of Tversky & Kahneman's (1973) research into probability judgements in inferential and statistical reasoning. Their theory suggested that people judge the probability of an event by the ease with which examples of the event can be recalled. Likewise, Pollard (1982) suggested that people might utilise availability when evaluating propositional problems. Essentially, the theory suggests people select answers that correspond to familiar information or familiar associations. This also implies that the more experience a person has with a set of outcomes, the more likely they are to accept the outcomes if they are valid, or reject them if they are false. Availability can be viewed with respect to *strength* and *activation* of memory items (Anderson, 1990; Baddeley, 1986). Items which are high in strength

[1] Although Evans (1995) asserts that these analytic processes exist, he does not say how they work.

and activation are more available to working memory. This says very little about how people reason on the abstract selection task (other than to say that people who are familiar with it will tend to get it right!). But it predicts that people will reason better on the thematic selection task when they have direct experience with the problem described, or are familiar with its content[1]. A good example of this is Griggs & Cox's (1982) "Drinking Age Rule" shown in section 1.4. The general rule of only being able to drink alcoholic beverages after a certain age is familiar to most adults in the western world, and these people tend to do very well on this version of the selection task. However, various experiments by Cosmides (1989) and Gigerenzer & Hug (1992)[2] show that availability cannot account for much of the facilitation on the thematic selection task. Both sets of experimenters compared selection task performance for "familiar descriptive" problems with Cosmides' "unfamiliar social contract" problems. They found that the social contract problems, although supposedly unfamiliar to the participants, produced much higher facilitation than the familiar descriptive problems. Ironically, Gigerenzer & Hug's (1992) second experiment exploring unilateral cheating options also showed a high level of **P & ~Q** responding (61%) for the a switched-social contract version of the "Drinking Age Rule" when their theory predicted only a low level of correct responding (*summarised in Table 3,* Gigerenzer & Hug, 1992, p.161). This high level was predicted by availability. Likewise, the availability heuristic correctly predicted a "middling to low" (48% and 38%) level of correct responding for "familiar descriptive" problems in experiments 1 and 2 of Cosmides' paper (*summarised in Table 2,* Cosmides, 1989, p.213). Therefore, although the availability heuristic cannot explain *all* of the facilitation on the thematic selection task, it may still have a significant effect on problems with familiar content.

Evans (1995) rejects the availability heuristic in favour of his own relevance heuristic. The relevance heuristic (Evans, 1984) suggests that people will choose answers that seem most *relevant* to the answering of the problem. For instance, in the "Drinking Age rule" the participant is cued to the perspective of a police officer – a person whose job is to check for *violations* of the law. Hence, the "drinks beer" and "16 year old" cards (**P & ~Q**) are seen as the most relevant to the given role. Evans suggests that the pragmatically relevant information "pops out" in the consciousness of the reasoner from the unconscious heuristic process in much the same way as useful chess moves "pop out" in the consciousness of an expert chess player[3] (Evans, 1995,

---

[1] Although this suggests that reasoning will be better in certain familiar domains, the availability heuristic is still considered a global theory because the same principle is presumed to govern reasoning within each of these domains. The same applies to relevance.

[2] Cosmides' and Gigerenzer & Hug's experiments will be reviewed in greater detail in Chapter 2.

[3] This analogy appears circular. Evans suggests that many researchers do not question the unconscious processing involved in deciding chess moves. He claims that this is validation for not needing to explain the unconscious processing involved in relevance. However, many others would argue that a primary aim of deductive research is to explicate implicit processes such as deciding a chess move!

p.150). There is some corollary evidence to back up the relevance heuristic. Evans conducted a selection task experiment that examined which cards were selected and how long each card was inspected. In both abstract and thematic versions of the task he found that only the cards participants finally selected were inspected for any significant length of time. Evans claims that this indicates participants had unconsciously pre-selected their responses (while reading the text) before inspecting the cards (Evans 1995).

The relevance heuristic has had some success in predicting the Matching bias on the abstract selection task and the content effect on thematic tasks (Evans, 1995; Evans et al., 1993). But the Heuristic-analytic theory suffers from much that same problem as the Mental Models approach – there is too much unexplained processing involved. In the case of Mental Models, the "fleshing out" process of the implicit model is left undefined; likewise, the processes involved in the analytic stage of the Heuristic-analytic theory are also not explicitly stated. Interestingly, Evans (1995) has made some attempts to combine the two theories. The heuristics mooted in the Heuristic-analytic theory could help explain the hitherto undefined processes involved in Johnson-Laird and Byrne's "fleshing out" of the implicit model, and the Evans' theory could benefit from utilising the construction of mental models in order to explain the processes in the "analytic" stage. I find this a promising line of research, however I also envisage that one or both theories will need to be reformulated. The Heuristic-analytic theory maintains that the unconscious heuristic selection occurs *before* any analytic processing, yet the Mental Model theory suggests that the "fleshing out" occurs *after* the formulation of the initial model and the drawing of a putative conclusion. This inconsistency may yet prove an unresolvable sticking point to the integration of these two theories.

The last theory reviewed in this chapter is a domain-specific theory known as *Pragmatic Reasoning Schemas* (Cheng & Holyoak, 1985). The underlying idea of this theory is that people reason best within certain thematic domains because they develop particular schema which direct their thinking within these domains. Schemas are specialised knowledge structures that encapsulate general structural information about groups of instances within a particular domain of experience. Schemas are built up empirically, being learnt or induced through direct experience of these given instances. This general structural information is fairly abstract in nature. The schema can be instantiated to particular cases by the filling in of 'slots' or 'variables' from the content of the real world instance (Anderson, 1990). The theory of Pragmatic Reasoning Schemas suggests that schemas can be used to reason-by-analogy from one thematic instance to another within the same domain. This is because there is a

core structural similarity between the two instances which can be mapped from the first instance to the next. Likewise, the structure of the answer to the first problem can be mapped onto the second problem.

A number of PRS (pragmatic reasoning schema) have been proposed; chief amongst them are the "*permission schema*" (Cheng & Holyoak, 1985) and the "*obligation schema*" (Cheng et al., 1986). The permission schema is built up from our experience of situations which require permission to perform an action. This schema encompasses four production rules:

(1) If the action is to be taken, then the precondition must be satisfied.
(2) If the action is not to be taken, then the precondition need not be satisfied.
(3) If the precondition is satisfied, then the action may be taken.
(4) *If the precondition is not satisfied, then the action must not be taken.*

The production rules with most relevance to the selection task are (1) and (4), which correspond to **P** (the action is taken) and **~Q** (the precondition is not satisfied) respectively. With respect to the "Drinking Age Rule", the 'action' relates to the drinking of alcohol and the 'precondition' is that you must be over a certain age. Therefore, the permission schema dictates that the cards "Drinking a Beer" (*the action is taken*) and "16 Years Old" (*the precondition is not satisfied*) must be turned over. The obligation schema has a similar structure, but a different context. It is built up from experience of situations which require an obligatory action once a specific precondition has been satisfied:

(1) If the precondition is satisfied, then the action must be taken.
(2) If the precondition is not satisfied, then the action need not be taken.
(3) If the action is taken, then the precondition may have been satisfied.
(4) *If the action is not taken, then the precondition must not have been satisfied.*

An example of this might be "If a golfer wins a round of golf, then he must buy a round of drinks for his playing partners". Permission and obligation schemas are often viewed as bilateral[1] opposites.

These schemas have had a lot of success in accounting for a number of phenomena related to the thematic version of the selection task. Foremost, of course, they were designed to account for the general level of facilitation found in these forms of thematic tasks, which they do fairly well. But they also correctly predict that a number of other thematic tasks will not provide reliable facilitation. For instance, neither Wason & Shapiro's "Towns and Transport" rule (*"Every time I go to Manchester, I travel by train"*) nor Manktelow & Evans' "Food and Drinks" rule (*"If I eat haddock then I drink*

---

[1] This means that different characters will have different viewpoints of the same problem. Take this example: (Mother to son) "If you do your chores, I'll give you an ice block". From the son's perspective this is a permission schema – he will get an ice block (action) once the chores are done (precondition). But from the mother's perspective it becomes an obligation schema – once the chores are done (precondition), she will have to give him an ice block (action). Bilateral reasoning has been examined quite extensively (Gigerenzer & Hug, 1992; Over & Manktelow, 1993; Liberman & Klar, 1996; Manktelow & Over, 1991), but a fuller review of the research is beyond the scope of this thesis.

gin") (in Evans et al., 1993) are expressed as permission or obligation schema, and they both produce levels of correct responding below 30%[1].

Another finding that pragmatic reasoning schema have had success with is the explanation of transfer effects. Traditionally, the selection task has been very resistant to transfer of facilitation from one task to the next. For instance, preceding an abstract task with the "Drinking Age" rule provides only minimal facilitation for the abstract task. But Griggs and Cox (1982) found that their "Clothing Age" rule (*"If a person is wearing blue, then that person must be over 19 years of age"*), which, when presented by itself, normally elicits only low levels of correct responding, was significantly facilitated when preceded by the "Drinking Age" rule. Evans et al. (1993) argue that participants do not see the "Clothing Age" rule as a permission rule unless the permission schema has already been elicited by the more familiar "Drinking Age" rule[2].

Girotto, Gilly, Blaye & Light (1989, in Evans et al., 1993) found more corollary evidence for the schema theory in their work with children. Firstly, they found 10-11 year-olds gave high rates of correct responding for permission versions of the selection task (which is at odds with Piagetian theory) but not for abstract selection tasks. But, more astonishingly, they also found that children as young as 6-7 years produced high levels of correct responding for permission versions of the "Reduced Array Selection Task"[3]; significantly more than for abstract versions.

Finally, there is another piece of research by Cheng & Holyoak (1985) which originally appeared to sit strongly in favour of their schema theory, but is now problematic. They constructed an *abstract* version of their permission schema:

If one is to take action "A", then one must first have fulfilled precondition "P".

Participants were cued to the position of an authority checking to see whether or not people were obeying the above regulation. They were then presented with four cards representing four different people: *"has taken the action"* (**P**), *"has not taken the action"* (**~P**), *"has fulfilled precondition P"* (**Q**), and *"has not fulfilled precondition P* (**~Q**). Cheng and Holyoak found that 61% of participants gave the correct **P&~Q** answer, compared to 19% for a similar non-permission-schema task using abstract letter and numbers. Hence, Cheng and Holyoak claimed that their schemas were content independent – the abstract structure of the schema was the facilitating factor, not the content of the

---

[1] It should be noted that a number of other theories (such as availability and relevance) also predict this.
[2] Again, this suggests that availability must be playing *some* part in this process, if only to trigger the permission schema.
[3] The RAST is identical to the standard version of the selection task except there are only two cards (**Q** and **~Q**) to choose from.

problem. However, Jackson & Griggs (1990) criticised their claim on methodological grounds. They noticed that Cheng & Holyoak had used explicit negatives ("has *not* taken the action", and "has *not* fulfilled precondition"). It has previously been found that the use of explicit negatives reduces matching bias and increases logical performance on truth table tasks (Evans et al., 1993). When Jackson and Griggs replicated Cheng & Holyoak's experiment using implicit negatives (such when the characters "took a *competing* action", or "filled a *competing* precondition"), they found the facilitation disappeared[1]. Since the permission schema should treat implicit and explicit negatives as semantically identical, this finding seems to invalidate the notion of the schema being content independent. In a related criticism, Cosmides (1989) pointed out that a number tasks which have the structure of a permission or obligation schema do not always produce facilitation. Take the "Clothing Age" rule again: although it has the definite structure of a permission rule, it does not seem to elicit facilitation on its own – it needs to be preceded by the "Drinking Age" rule. This failure of content independence casts an undeniable shadow-of-doubt on the validity of pragmatic reasoning schemas – at least, in their current form.

The schema theory has also been attacked by another domain-specific theory, Leda Cosmides' (1989) Social Contract theory. This theory suggests that the facilitation found in social and deontic versions of the selection task can be explained in terms of evolutionary theory rather than induced cognitive structures such as pragmatic reasoning schema. Cosmides claims that pragmatic reasoning schema are neither necessary nor sufficient to facilitate correct reasoning on the Wason Selection Task. The next chapter will examine her claims more closely and critique the theory that she espouses: The Theory of Social Exchange.

---

[1] However, using explicit negatives in abstract "number and letter" selection tasks does *not* facilitate level of correct responding, although it does seem to reduce *matching bias*.

2 Cheater Detection, Social Exchange and Task Understanding

2.1 Cheater Detection

Cosmides theory starts from a uniquely different direction to that of all other current theories of reasoning. Instead of noting a phenomenon then constructing a theory to account for it, she utilised Marr's (1982) computational approach to construct her theory from base principles. A computational theory starts by asking "What must happen if a particular function is to be accomplished". Marr argued that to understand the mechanisms which govern a function, we must first understand the problems that these mechanisms were 'designed' to solve. While looking for mechanisms to account for reasoning phenomena, Cosmides examined the evolutionary history of humans and followed a series of assumptions through to the construction of an evolution mechanism. Her idea was to take what we know about the environmental conditions which existed in the Pleistocene era (under which Homo sapiens have undergone 99% of their evolutionary history) and posit the sort of selection[1] pressures that our ancestors must have had to deal with. Then she posited mechanisms that may have evolved to deal with these selection pressures. Finally, she looked for evidence of these mechanisms at work today as validation that these mechanisms exist.

In the first step Cosmides draws our attention to the fact that Pleistocene hominids lived in small social groups or communities (Wilson & Sober, 1994). They evolved as socially gregarious creatures, preferring the added benefits to *inclusive fitness*[2] afforded by *reciprocal altruism*. Altruistic behaviour is behaviour exhibited by one individual which benefits the inclusive fitness of another organism while being detrimental to the inclusive fitness of the first individual (Trivers, 1971). True altruistic behaviour has been shown to be an evolutionary 'dead end' because they are of no evolutionary benefit to the individual who exhibits the behaviour, indeed such behaviour is assumed to be selected against (Hamilton, 1964). But two forms of "altruistic behaviour" exist which do not suffer this fate. Firstly, there is *kin selection*, where the individual 'helps' an offspring, sibling or another close relative (Hamilton, 1964). In such a case, the act is no longer considered detrimental to inclusive fitness because the individual is 'helping' another organism who shares a good deal of the same genes. By

---

[1] In this case and in the following few paragraphs "selection" refers to *Natural Selection* (in the evolutionary sense), not to the Wason Selection Task.

[2] "Inclusive Fitness" is Hamilton's (1964) expression used to signify an individual's likelihood of successfully passing on its genes to the next generation.

raising the inclusive fitness of a close relative you are in effect increasing the chance of your own genes being passed on to successive generations. Hence, such behaviour is actually selected for. The other evolutionarily valid form of altruism is *reciprocal altruism* (Trivers, 1971). This is the case when one organism benefits another organism's inclusive fitness, to the detriment of their own, on the assumption that the other organism will reciprocate the benefit at a later time. In essence, it is the notion of "doing the other individual a favour". Four conditions must be fulfilled in order for a species to develop reciprocal altruism:

(1) The members of the species must interact often enough with whatever organisms[1] they are to reciprocate with to ensure that a net benefit amasses to all donors.
(2) The benefit of receiving aid must exceed, on average, the cost of donating.
(3) Any dominance hierarchies must, to a certain degree, be flexible.
(4) Individuals must be able to recognise previous recipients who have failed to reciprocate and withhold further benefits.

It is the forth condition which Cosmides saw as most relevant in the conception of her theory. She viewed detection of cheaters as fundamental to the evolutionary development of reciprocal altruism – or, in the human context, *Social Exchange*. She suggested that when two individuals engaged in social exchange they entered into a mutual agreement whereby one party *supplied a benefit* to the second party on the understanding that the second party would *pay a cost* to the first party. Cosmides called this cost-benefit structure a *Social Contract*. If an individual could not reliably detect others who have a history of cheating on social contracts, then they would be in danger of being constantly cheated. This would be at the detriment of that person's inclusive fitness. When extrapolated across the whole human species, we would expect that reciprocal altruism could not remain evolutionarily stable and would soon be selected out.

To accomplish this cheater detection, Cosmides claims that humans have evolved a *cheater-detection algorithm* which incorporates the cost/benefit structure of a social contract. This cheater-detection algorithm is one of a posited set of genetic mechanisms known as *"Darwinian algorithms"* (Cosmides & Tooby, 1987; Cosmides, 1989). Darwinian algorithms are "*specialized learning mechanisms that organize experience into adaptively meaningful schemes or frames*" (Cosmides & Tooby, 1987, p286). These algorithms are triggered by environmental or proprioceptive domain-specific information. When activated, they channel attention, perception and memory, and recall specialised procedural knowledge that lead to domain-appropriate

[1] Normally this will be another member of the same species, but a number of instances of inter-species reciprocation have been found (for instance, the cleaner-fish's symbiotic relationship with it's "host").

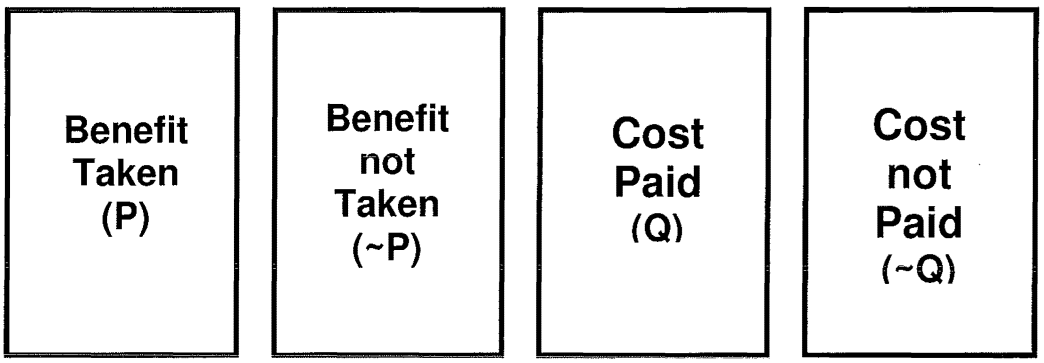
inferences, judgements and choices. In a sense, they are like the cognitive schemas that underlie pragmatic reasoning schema except they have a genetic origin – not simply a learnt one - and are evolutionarily adaptive.

2.2 Social Contracts

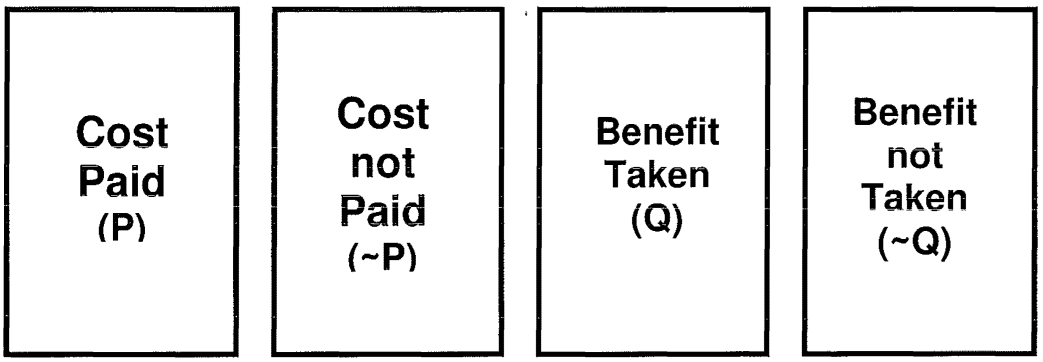
The critical feature of Cosmides' theory is the cost/benefit structure of the social contract. She proposes two versions of this structure:

(1) **Standard Social Contract:** *"If you take the benefit, then you pay the cost".*

When applied to the Wason Selection task, this leads to the following cards:



And (2) **Switched Social Contract:** *"If you pay the cost, then you take the benefit".*



The cheater-detection algorithm should lead one to:

(1) Watch out for an individual if they have accepted the benefit.

(2) Watch out for an individual if they have not paid the cost.

(3) Ignore an individual who has not accepted the benefit.

(4) Ignore an individual who has paid the cost.

In both cases Cosmides predicts that the cheater-detection algorithm should alert the individual to the possibility of the *benefit being taken without the cost being paid.* This leads to an interesting prediction. While the cheater-detection algorithm should induce people to select the **P&~Q** cards for the standard social contract, it should also

induce people to select the **~P&Q** cards on the switched social contract. Previous to Cosmides' theory, the selection of **~P&Q** cards was a rare response finding in the literature (Evans et. al., 1993).

Cosmides claims that her theory is not restricted to one-on-one social contracts (*personal exchanges*) but also includes contracts imposed on members of a community as *social laws*. For instance, the "Drinking Age rule" is presented in the context of a person whose job is to check whether people are breaking a social law. In such cases the notion of "cost" often becomes blurred. Take the "Drinking Age rule" again: the act of "drinking beer" can certainly be considered a benefit, but it is hard to imagine the condition of "being over 19" as a cost. Because of this, Cosmides suggests that the notion of "preconditions" can be substituted for costs. This weakens her position considerably in my view, because such a concession seems closer to Pragmatic Reasoning schema. Indeed, many researchers have claimed that social contracts are nothing more than specific instances of permission and obligation schema (Cheng & Holyoak, 1989; Politzer & Nguyen-Xuan; 1992).

Anticipating this criticism, Cosmides set out to test her theory against the predictions of Pragmatic Reasoning schema and availability theories. She had three basic hypotheses:

(1) Availability is neither necessary nor sufficient to induce facilitation of correct responding on the Wason Selection Task, while even completely foreign cases of social exchange would produce high levels of facilitation.

(2) Pragmatic Reasoning Schemas only produce facilitation in cases where they conform to the boundaries of social contracts.

(3) Problems that conform to Switch Social Contracts will produce high levels of **~P&Q** responses (availability and PRS theories do not predict this).

2.3 Evidence for Social Exchange theory

In her first set of experiments Cosmides set out to test her first and third hypotheses (Cosmides, 1989). She started with four conditions – *unfamiliar (standard) social contract tasks*, *unfamiliar descriptive tasks*, *familiar descriptive tasks* and *thematic-abstract tasks*[1]. The unfamiliar social contracts tasks conformed to the structure of a social contract incorporating a cost/benefit relationship and prompting cheater-detection, but the content was culturally unfamiliar to the participants. The unfamiliar

[1] Cosmides (1989) refers to these as just "abstract problems". But they are only *abstract* in so much as their underlying structure is of an abstract design. They do not adhere to a format or structure that any theory of reasoning predicts should produce facilitation. They are, however, framed in a thematic content scenario. Hence I refer to them now, and for the rest of this thesis, as *thematic-abstract* tasks. This distinction is necessary because later other scenarios will be presented which "mix" the content component and structural components (such as "abstract-social-contracts" and "pure-abstract" tasks).

descriptive tasks contained the same content as the social contract version, but did not incorporate the cost/benefit relationship or elicit cheater-detection. The familiar descriptive task contained content derived from every-day tasks that were familiar to the participants, but did not contain the cost-benefit relationship or cheater-detection. Finally, the thematic-abstract tasks were rules of an abstract structure with only nominal thematic content. (All of these tasks are presented in Appendix A1.1 – A1.4). Cosmides used at least two versions of each task; subjects received one version of each condition. Because social contracts can be in the form of a *social law* (agreed upon by an entire population) or a *private exchange* (an agreement between two parties), Cosmides included both types in her unfamiliar social contracts conditions.

The availability hypothesis predicts that unfamiliar social contract, unfamiliar descriptive and thematic-abstract tasks should elicit only low levels of correct responding (10-25%), while familiar descriptive tasks should provide higher levels of correct responding (say 40-50%). The Social Exchange theory, on the other hand, predicts that the unfamiliar social contracts should elicit much higher levels of correct responding (above 60%). Cosmides' results appeared to confirm this: the unfamiliar social contracts conditions elicited 73%[1] correct responding, significantly more than unfamiliar descriptive tasks (23%), thematic-abstract tasks (27%) and familiar descriptive tasks (42%). Hence, Cosmides concluded that availability by familiarity is neither necessary nor sufficient to elicit facilitation on the selection task, while social contracts did appear to produce facilitation.

Further evidence was found in Cosmides next set of experiments. This time she tested her prediction that switched social contracts should elicit high levels of **~P&Q** responses. As mentioned earlier, this response was only very rarely found prior to Cosmides' experiments. She used the same tasks as used in her first set of experiments, except that the rules were 'switched' in the unfamiliar tasks[1]. The availability hypothesis predicted that instances of **~P&Q** should be very low across all conditions, while Social Exchange theory predicted high levels of **~P&Q** responses only in the switched social contract conditions. Again, this is precisely what Cosmides found. The unfamiliar switched social contracts elicited the **~P&Q** response 71% of the time, while the other conditions averaged 0-2%.

Cosmides next step was to test Social Exchange theory against Pragmatic Reasoning Schema. She tested these theories by examining performance on *social contracts* and *non-social-contract permission* tasks. In the social contracts tasks, the

---

[1] Each of the figures presented here are the means across two experiment of the same design. The same is true of the other results cited from Cosmides' paper.

surrounding scenario framed the rule being tested with a clear cost/benefit relationship and cued the participant to the role of cheater-detection. The non-social-contract permission tasks used precisely the same wording of the tested rule, but the surrounding scenario did not frame the rule in a cost/benefit relationship (rather, it was framed in an "action"/"precondition" relationship) and no cheating was implied. Cosmides also broke these conditions down into culturally familiar tasks (the "Grover High School" tasks) and culturally unfamiliar tasks (versions of the "Cassava Root" and "Duiker Meat" problems). (*For the full text of each task see Appendix A1.5 & A1.6*).

Pragmatic Reasoning schema predicts that there should be no difference in the level of correct responses (**P&~Q**) because it views all the tasks as permission schema, while Social Exchange theory predicts an increased level of performance for the social contracts tasks over the non-social-contract permission tasks. The results again seemed to favour Social Exchange theory. While social contracts tasks elicited 77% correct responses, the non-social-contract permission tasks elicited only 37% correct responses.

Cosmides then repeated the study using switched social contracts and corresponding "switched non-social contract permission rules". These switched permission rules take the form "*If the precondition is satisfied, then the action may be taken*". In such a case, Cosmides says pragmatic reasoning schemas predict that the "action is taken" card (**Q**) and the "precondition is not met" card (**~P**) should be selected (as is the case for social contracts). Social Exchange theory, of course, predicts that only the true switched social contracts tasks will elicit a high level of **~P&Q** responses. The scenarios used were the same as used in the preceding social contract/permission schema study except the rules were switched. Yet again, Cosmides predictions were correct; 73% of the switched social contract responses were **~P&Q**, compared with only 5% for the switched non-social-contract permission rules. This experiment, and the previous one, indicate that people seem to be guided by social contracts rather than pragmatic reasoning schemas. At the very least, these studies seem to show that social contracts are not simply specific cases of permission or obligation schema.

Cosmides' theory has since been refined by Gigerenzer & Hug (1992). Their first experiment attempted to disentangle the cost/benefit relationship and the cheater-detection algorithm in the social contract rule. They produced two types of social contracts scenarios which each contained the same tested rule. In one condition the participant was cued to the perspective of someone charged with enforcing a social

---

[1] It is interesting to note that Cosmides never indicates whether the propositions were switched for her familiar descriptive tasks; one presumes not. If not, then this is creates problems comparing results from her familiar and unfamiliar conditions!

contract rule; in the other condition the participant was cued to the perspective of someone interested in determining whether a certain social contract rule is being applied by the appropriate authorities. The aim of the first condition is cheater-detection, but the second condition does not require cheater-detection even though the same cost/benefit relationship is evident in the tested rule. Corresponding examples of these two types[1] are given in *Appendix A1.7.* Gigerenzer and Hug found that the cheater-detection condition elicited a correct response 83% of the time, while the no-cheating condition elicited only a correct response 45% of the time.

They concluded that the facilitating factor in social contract tasks was not the cost/benefit relationship of the contract per se, but the *search for cheaters.* However, their experiment did not completely disentangle the cost/benefit relationship from the cheater-detection component. Their experiment only showed that cheater-detection seems necessary to induce facilitation. They did not examine whether cheater-detection could induce facilitation without the presence of a cost/benefit relationship imbedded in the tested rule. This possibility is the primary focus of the first set of experiments, which are presented in the next chapter. Before this, though, criticisms of Social Exchange theory are reviewed along with other posited alternatives.

2.4 Criticism of Social Contracts

The first criticism of Cosmides' experiments is perhaps rather a trivial one. Evans et al. (1993) suggest that since the mid 1980's no one has seriously considered availability to be the leading factor contributing to high facilitation on the Wason selection task. Hence, they claim, Cosmides' experiments testing social contracts against availability are essentially 'straw arguments'. This may be true to a point, but I believe that they were important for the following reasons. They showed that social contracts facilitate performance even in the absence of experience of the given context, and they showed the previously unheralded effect of high levels of **~P&Q** responding under "switched" conditions. Beyond this, although the availability hypothesis was essentially a "dead issue" by the time Cosmides' paper finally was published in 1989, this was not the case when these experiments were originally carried out for her PhD, completed in 1985. The general point that Evans and colleges make is still a reasonable one, however. Hence, further attention is focused on the criticisms of switched social contracts and to the comparisons of Social Exchange and Pragmatic Reasoning Schema.

---

[1] Four different scenario contexts were used. I have only reproduced one permission rule here, for the sake of brevity. The figures, however, are a mean of the four scenarios across each condition.

Cheng & Holyoak (1989) began the attack on Social Exchange theory. Among a number of criticisms were two particularly compelling points. First of all, they pointed out the existence of a number of thematic tasks that induce reliably high levels of facilitation, yet do not fit the form of a social contract. One of these is the well-documented "Sears department store rule" (D'Andrade, 1982; in Cheng & Holyoak, 1989). In this scenario the participant is cued to the perspective of a store attendant at Sears department store. The rule is "*If any purchase exceeds $30, the receipt must have the signature of the department manager on the back*". Participants reliably indicate to turn over cards corresponding to "*purchase exceeds $30*" (**P**) and "*no signature on back of receipt*" (**~Q**). Yet, as Cheng and Holyoak point out, having a purchase exceed a certain value can hardly be seen as a benefit to either the purchaser or the store attendant (why not make the prerequisite $10, or $100?), and the act of signing the back of the receipt is certainly not a cost or precondition. It does, though, satisfy the form of an obligation schema. Strangely, Cosmides claimed that the "Sears problem" *is* a social contract (Cosmides, 1989, p200), but she failed to explain why! A second rule which induces reliable facilitation, but does not fit the benefit/cost relationship is "*If you clear up spilt blood, you must wear rubber gloves*" (Manktelow & Over, 1991). Again, it is hard to see how "clearing up split blood" is a "benefit", although "wearing rubber gloves", while certainly not a genuine "cost" does appear to be a precondition. Yet it still seems to satisfy the form of a permission schema much more easily than a social contract.

The second compelling point that Cheng and Holyoak made was that Cosmides' rules appear to be *biconditional.* Politzer & Nguyen-Xuan (1992) extend this criticism. They claim that the very nature of a social exchange ("*If you give me P, I will give you Q*") necessitates a biconditional rule. The fact that both parties agree to the contract at all indicates that each expects the other party to give them what they want. For one party to agree to the contract, but not care whether they get what they want, would be simply foolish!

This begs another question, though. If Cosmides tasks are biconditional, then why do the participants tend to give **P&~Q** responses (or **~P&Q** for switched scenarios). It has long been established that the correct *logical* response to a biconditional rule on the Wason Selection Task is to select *all* the cards (**P,~P,Q,&~Q**). Yet Cosmides reports virtually no instances of this in her data. The reason, Politzer and Nguyen-Xuan claim, is because Cosmides' detailed scenarios place extreme emphasis on checking for violations in only one party's side of each contract. This hypothesis seems, at first glance, to be an easy one to test: an experiment could be devised whereby the participant is cued to the perspective of somebody charged with checking whether

either of two parties is cheating on a social contract. In such a case, Politzer and Nguyen-Xuan would predict a high number of cases where participants (correctly) select all four cards. Unfortunately, so would Cosmides! Social Exchange theory would predict that the cheater-detection algorithm would be triggered twice, once for each party. Hence, such an experiment is incapable of testing between Politzer and Nguyen-Xuan's hypothesis and Social Exchange theory.

The biconditional criticism expressed by Cheng & Holyoak and Politzer & Nguyen-Xuan is note-worthy with regards to a number of social contracts, but I believe that it does not apply in all cases. The situations that it does seem to apply to are the ones Cosmides term *"private exchanges"* (see Appendix A1.1). In these scenarios the two parties concerned agree to share resources in a reciprocal "deal". In theory, *either* party could cheat the other party, *although in practice the scenario orders the temporal sequence in such a way that only one party can cheat.* For instance, in the "*Duiker meat – private exchange*" scenario Bo, the "crafty old Namka", ensures that he gets his ostrich eggshell *the evening before* he claims he will supply his guests with duiker meat. Likewise, on the "*Cassava root – private exchange*" scenario Big Kiku ensures that his guests get the tattoo on their face *before* he claims he will give them the cassava root. Therefore, in both Cosmides' "private exchange" cases we have an *initial* biconditional contract between the characters, but the *final* rule being tested in still uni-conditional. Hence, I do not see the fact that private social contracts seem to be biconditional as necessarily a problem for Social Exchange theory – the facilitation found in these tasks could still be due to the cheater-detection algorithm searching for cheating in one "direction" across the biconditional contract. Of course, this does not rule out the possibility that Politzer & Nguyen-Xuan's criticism may be valid either; I suggest that the only conclusion that can be drawn from this debate is that there is not enough evidence to draw a conclusion.

Although the "private exchange" social contracts are biconditional, I am not convinced the same can be said of Cosmides' "social law" social contracts. In these cases, the rule is a law *imposed* on a society's members by its relevant leaders or lawmakers. The members do not have a choice regarding acceptance of the contract. Although adherence to the rule may benefit the society as a whole, it does not appear to benefit the individual adhering to the rule. He or she still has to pay the cost or fulfil the precondition in order to receive the benefit. It would be much more efficient, from the individual's point of view, to take the benefit regardless for whether they have paid the cost or not. For instance, in the case of the "Cassava root" problem, if the rule was not in place, the unmarried individual could take as much cassava root as they

pleased, so the enforcement of the rule is an impediment to them. But, the rule is beneficial to the Kaluame society because it rations the usage of a valuable resource. Hence, the social law, while beneficial for the society as a whole, is simply an unnecessary barrier to the benefit from the point of view of the individual. On the other hand, these social laws can only realistically be cheated on by the individual – it would be silly for the society to cheat the individual by denying the individual the benefit after he has paid the cost (or fulfilled the precondition). For instance under the "drinking age" rule, the possibility of a public bar denying an individual an alcoholic drink after the individual has reached the required age appears ludicrous (and particularly bad for business!).

Politzer & Nguyen-Xuan (1992) further criticised Cosmides for her interpretation of "switched" conditions. Cosmides claims that only social contracts theory predicts that switched conditions will elicit the **~P&Q** responses typically found in these cases. Politzer and Nguyen-Xuan say that this is false – in fact, any semantically based reasoning theory would predict this outcome. They even go as far as to claim that "mental logic" theories (Formal Rules of Inference) would make this prediction, as *"the rules of logic do not apply to an uninterpreted surface structure, but, rather, to a propositional form that is the product of semantic and pragmatic processing"* (1992, p406). It is still questionable whether Formal Rules theory would really predict the high levels of responding that Cosmides found (remember, it does not even predict the general "content effect"), but Politzer and Nguyen-Xuan have made a valuable point. That is, Cosmides claims that other theories do not "see though" the *syntax* of her switched rules, but this is not necessarily true. Johnson-Laird & Byrne (1991) make a similar point. They claim that participants may reinterpret the switched social contract form as a standard social contract. For example, Johnson-Laird & Byrne suggest that participants take *"If a man has a tattoo on his face, then he eats cassava root"* to mean *"A man may eat cassava root only if he has a tattoo on his face"*. Unfortunately, Johnson-Laird & Byrne give no reason *why* participants would spontaneously interpret the rule this way.

Here is one suggestion. It seems evident that Cosmides switched "private exchange" social contracts are *temporally inconsistent*. Take her switched rules for the "Cassava Root" problem (*"If I give you cassava root, then you must get a tattoo on your face"*) and her "Duiker Meat" problem (*"If I give you duiker meat, then you must give me your ostrich eggshell"*). This implies that the person proposing the contract supplies a benefit to the second party either *before*, or *at the same time* as, the second party pays the cost to the 'proposer'. But, in the text of the scenarios, the potential cheater (the

'proposer') always makes sure *he receives his benefit first* (Cosmides, 1989, p217). In the "Cassava root" problem, the tattoo had to be in place *before* the cassava was given (temporally, not contractually). Likewise, in the "Duiker meat" problem, the ostrich eggshell had to be presented *before* the duiker meat was given. Hence, a temporal inconsistency occurs between the implication of the tested rules and the series of events in the main text of the scenarios. The only way to let the text of the scenarios fit with the tested rules is to imply that the costs were paid *before* the benefits were supplied. Therefore, the participants would have had to interpret the rules to read *"If you get a tattoo on your face, I will give you cassava root"* and *"If you give me your ostrich eggshell, I will give you duiker meat"*, which are, of course, the *standard* versions of the social contracts.

Cosmides switched "social law" social contracts do not suffer this same temporal problem. Yet they still induce the same high rates of **~P&Q** responses as the "private exchange" scenarios. The only foreseeable explanation for this is Politzer & Nguyen-Xuan's notion that the rule is biconditional. Although there are reservations about the validity of this notion, it does explain why the propositions in the "social law" rules can be successfully reversed. Note that in a causal rule such as "*If the ball rolls right, then the green light will come on*", switching the propositions will switch the chain of causal events ("*If the green light comes on, then the ball will roll right*"). In the case of the "social law" social contracts, switching the propositions does not change any causal contingencies - the content of the scenario explicitly states the preconditions must be satisfied *before* the benefit may be taken. This is still the case whether the rule is standard or switched. Hence, in both forms, the rule is semantically equivalent.

This "biconditionality" hypothesis seems compelling, but it is unable to be empirically distinguished from Social Exchange theory. As mentioned earlier, even if social contracts are biconditional, Social Exchange theory could claim that this just sets up *two* cheater-detection algorithms – one for each party.

2.5 A global theory of task understanding:

A more damaging attack on Social Exchange has been presented by Liberman & Klar (1996). Instead of appealing to domain-specific theories of the content effect, such as Social Exchange and Pragmatic Reasoning Schemas, they have forwarded another global theory. The theory states that a person's performance on the Wason Selection task is determined by their understanding of what the task entails. The crux of the theory is that people *do* tend to reason in a logical and rational manner given what they understand about the task. There are *three fundament aspects* that Liberman and Klar suggest influence task understanding:

(1)	***The clarity of the rule in terms of determination and direction.***
(2)	***The nature of the alternative to the tested rule and the falsifying instance it entails.***
(3)	***The perceived relevance of looking for violations of the rule.***

In fact, these factors are not new ideas. All have been proposed, in one form or another, by other researchers. But Liberman and Klar were the first to take these ideas and formulate them into a coherent theory. Each of the aspects will be reviewed in greater detail.

The clarity of the rule in terms of determination and direction:

In order to reason in the way prescribed by formal logic participants have to understand that the rule is both *deterministic* and *uni-directional*. By "deterministic", Liberman and Klar mean that application of the rule will result in the same out come every time, without fail. They suggest that many selection task rules produce only low levels of facilitation because they fail to convey this. Instead, they provide or imply a probabilistic interpretation of the rule. Take this example: *"If an individual comes into contact with a sick person, then that individual will become sick too"*. Formal logic would claim that this rule is totally deterministic, yet we know from personal experience that contact with a sick personal does not necessarily result in one's self becoming sick. You might not catch the infection through immunity or luck, or you might only become a 'carrier' without producing symptoms of sickness. In such a case, it is reasonable to assume that the "rule" is probabilistic. In probabilistic cases, it is not possible to falsify the rule: finding an instance where a person *has* come into contact with a sick person, yet does not become sick does not mean that the rule is false in the probabilistic sense.

The participant also needs to understand that the rule is "unidirectional". By this, Liberman and Klar mean that the rule cannot be reversed (see Chapter 1.2). The participant needs to understand that "A implies B" does not necessarily mean that "B also implies A". As we have seen previously, people often interpret abstract rules as biconditional and, therefore, make errors on the selection task. But in the previous section we saw that biconditional rules may still elicit correct responses on thematic selection tasks if the surrounding scenarios place emphasis on evaluation of the rule from a given uni-directional perspective. In such cases, the rule could be thought of as "pseudo-directional".

The nature of the alternative to the tested rule.

Formal logic states that there is only one falsifying alternative to **"IF P⇒Q"** which is the instance of **P&~Q**. Liberman and Klar suggest that for reasoners to make selections in accordance with formal logic, they have to perceive this violating instance and not be confused with other (invalid) alternatives[1]. Therefore, performance should be better on selection tasks that have salient **P&~Q** alternatives. Likewise, performance will be worse on tasks that have salient alternatives *other* than **P&~Q**. For example: *"If I eat apples, I will be sick"* could have several salient invalid alternatives, depending on the nature of the preceding scenario. For instance, the person might be feeling sick because they have already eaten far too much. In this case *any* food eaten would make the person sick, so the violating instances would be **P&~Q** and **~P&~Q**. This, in turn, may lead to participants selecting **P, ~P,** and **~Q** on the selection task.

The perceived relevance of looking for violations of the rule.

The only way to logically test a rule is, of course, to look for instances that *falsify* the rule. But in everyday life falsification may not necessarily be the best policy. Liberman and Klar suggest that people will look for violations of a rule only if they perceive it to be *relevant*[2] to the completion of the task. They claim three factors contribute to this:

(a) *The perceived efficiency of negative vs. positive test strategies:*

In many situations a *positive* test strategy (looking for instances of *confirmation*) may be perceived as more efficient than a negative test strategy (looking for instances of violation) (Klayman & Ha, 1987). This might occur when participants perceive violations as impossible or when the perceived chance of finding a violation is low compared to that of confirmation. This is particularly true if there is no indication beforehand whether the rule being tested has any confirmatory evidence. Kirby (1994a, 1994b) demonstrated a similar effect. He found that the selection of each card on the selection task was related to the cards perceived "utility". Utility is related to the estimated probability of cards' ability to falsify or confirm the rule. For instance, take the conditional "*If it is a raven, then it is black*". It would be much more efficient to search the set of "ravens" looking for one that is "not black" than it would to search the set of "non-black things" looking for one that is not a "raven" (Over & Evans, 1994). Hence, the perceived utility of the "ravens" card is higher than that of the "non-black" card. Likewise, searching the "black" card may be perceived as worth while by the participant if they expect a high chance of finding a "raven' on the back.

---

[1] This appears to be the kind of process that Johnson-Laird & Byrne's (1991) Mental Model theory explains well.

[2] Their notion of relevance seems to have strong connections to Evans' (1984, 1995) heuristic, but they fail to cite him.

(b) *The activation of a "Detective set":*

It has been well established that if the explicit goal of a selection task is to look for violations of a rule, then the content of the accompanying scenario may induce facilitation. Van Duyne (1974; in Cosmides, 1989) found that facilitation increased if subjects were cued to look for falsification with what he called "detective sets". A "detective set" is simply a passage that cues the participant to the perspective of someone who looks for violations of the rule, such as the policeman in the "Drinking Age" rule. However, it has since been found that "detective sets" do not facilitate reasoning by themselves, they only appear to enhance performance on tasks that already produce facilitating effects (Liberman & Klar, 1996). Liberman and Klar claim that this effect incorporates nicely into their model: "[detective sets] *enhance* ***P&~Q*** *selection only once* ***P&~Q*** *is identified as the only relevant violating case.*" (Liberman & Klar, 1996, p132). I think that they might have their chain of cause and effect around the wrong way, here. I suggest that the activation of the detective set focuses the participant's attention to the relevance of falsification. Hence, participants then perceive **P&~Q** as the relevant violating selection.

(c) *The "Maturity" of the hypothesis*:

Liberman and Klar claim, for falsification to be perceived as relevant there must first be enough evidence to suggest that the rule is, or might be true. All too often in real life the option to falsify a given rule may seem bizarre, even though it is *logically* necessary. Take the example of a doctor who is to test a hypothesis connecting a suspected case of food poisoning. The hypothesis is "*If a person eats spoiled food, they will become ill*". While it may seem wise to test people who have eaten spoiled food (the **P** card) for instances of people who have not become sick, it seems ludicrous to begin looking *at healthy people* (the **~Q** card) for instances of people who have eaten the spoiled food. This is particularly true if the doctor does not have enough confirmatory evidence to suggest that the hypothesis might be true. Falsification only becomes relevant once the hypothesis or relationship is considered established.

Liberman & Klar set about testing their theory of task understanding against Social Exchange theory. Firstly, they took Gigerenzer & Hug's (1992) "cheating" and "no-cheating" tasks and systematically went through the text of each one to see which aspects of task understanding they successfully conveyed. The details were different for each scenario, but the following features tended to show up:

- Most of the *no-cheating* conditions suggested non-deterministic or biconditional rules, while the *cheating* conditions suggested deterministic, uni-directional rules (or at least "pseudo-directional" rules).

- Many of the *no-cheating* conditions suggested other alternatives to the **P&~Q** violating instances, while the *cheating* conditions did not.
- The *cheating* conditions (by their nature) cued the participant to a "detective set" and explicitly asked the participant to look for instances of violations. The *no-cheating* conditions cued the participant to the perspective of a "wondering visitor" who is unsure whether a social contract is being used or not (this suggests confirmation strategies may be useful as well as falsification).

Given these discrepancies between the cheating and no-cheating versions of these social contracts, Liberman & Klar decided to design corresponding versions which "unconfounded" each of the problems. The *unconfounded cheating* condition took the cheating tasks and added, subtracted or changed certain elements of the scenarios in ways that should *inhibit* selection of **P&~Q**. The *unconfounded no-cheating* condition took the no-cheating tasks and added, subtracted or changed certain elements of the scenarios in ways that should *facilitate* selection of **P&~Q**. (*An example of these is given in Appendix A1.8*).

Liberman & Klar then tested each of these unconfounded conditions against the original versions. They predicted that the unconfounded cheating versions would produce much less facilitation than the original cheating version, while the unconfounded no-cheating versions would produce much more facilitation than the original no-cheating versions. This is precisely what they found. Across four different thematic scenarios the original cheating condition produced 72% correct responses compared to only 31% for the unconfounded cheating condition, and the original no-cheating condition produced 35% correct responses compared to 70% for the unconfounded no-cheating condition. Together, these trends appear to confirm Liberman & Klar's theory, while casting considerable doubt on the need for a cheater-detection algorithm.

Yet Liberman and Klar's experiment may have methodological flaws. It is possible that changing the wording of the original cheating and no-cheating versions may implicitly destroy or create social contract rules. Take, for example, the unconfounded versions of the "Overnight Problem" (*see Appendix 1.8*). In the unconfounded cheating version Liberman and Klar say *"This version added to the original cheating version that according to the hiker's tradition, a hiker may deviate from the rule if he goes for an especially long track on the day after."* (Liberman & Klar, 1996, p.151). This appears to destroy the deterministic nature of a social contract (in fact, this is what Liberman and Klar were trying to achieve). By letting a person break the rule on the condition that they fulfil a second precondition, you have, in effect, two rules operating. Not only does this confuse matters considerably, but there is no way of discovering whether the first

rule has been broken until you know the outcome of the second rule as well. To make matters worse, Liberman and Klar say, *"In this case, he or she has to ask another hiker, who didn't sleep in the cabin, to bring wood."* This clearly takes the onus off the first hiker to conform to the rule at all. Overall, I find it hard to see how the rule can still be considered a true social contract. The rule may still have costs and benefits imbedded within it, but the structure of their relationship is blurred. I think Cosmides, and Gigerenzer & Hug, would claim that such a rule could not trigger the cheater-detection algorithm. Similar problems appear in Liberman & Klar's other unconfounded cheating options, although this is hard to establish for sure – we do not have the *actual* text used in the scenario for analysis[1].

However, this criticism does not extend to Liberman & Klar's unconfounded no-cheating version. As far as I can see, there appears to be no prompting of the participant to look for cheaters, yet these tasks were answered as accurately as the original cheating versions. At the very least, this indicates that cheater-detection is not *necessary* to induce facilitation on the Wason selection task (although it still may be *sufficient*).

2.6 Scope for further research:

There are two areas from the literature on Cosmides Social Exchange theory and Liberman & Klar's Theory of Task Understanding that I suggest warrant further investigation.

The first question is that of the necessity of the cost/benefit relationship of the social contract in Social Exchange theory. Gigerenzer & Hug (1992) showed that it was the cheater-detection algorithm that appeared to facilitate reasoning, but they could not disentangle the search for cheaters from the presence of the cost/benefit relationship of the social exchange. In chapter 3, I aim to create deceit-detection scenarios, which do not contain a cost/benefit relationship in the tested rule, and compare their performance against social contract scenarios. As a side issue, I will endeavour to construct these scenarios to the prescriptions which Liberman & Klar suggest should lead to good task understanding and, therefore, good task performance.

The second point examines one of the predictions of Liberman & Klar's theory. They suggest that good task performance derives from good task understanding – not domain-specific knowledge structures such as social contracts or pragmatic reasoning schemas. If this is the case, then even abstract tasks should produce high levels of facilitation if the surrounding scenarios can convey the important factors which

---

[1] This is a problem for critiquing all of Liberman & Klar's scenarios. One does not know the exact phrasing of each problem.

Liberman & Klar suggest lead to good task understanding. The aim of chapter 4 is to construct such tasks and compare their performance to that of social contracts.

3 Experiment One: Deceit Detection vs. Social Contracts

3.0 Introduction:

Previously, Gigerenzer & Hug (1992) managed to construct social contract scenarios that did not cue the participants to look for cheaters. They found that these scenarios produced significantly lower levels of facilitation than scenarios that required cheater-detection. From this, they concluded that it was the cheater-detection algorithm that facilitated reasoning. The problem with this is that while their experiments appeared to show that cheater-detection was *necessary* for facilitation[1], they did not show whether cheater-detection was *sufficient* to produce facilitation. This was in part due to the fact that Cosmides' original conception of the cheater-detection algorithm imbedded the cost/benefit structure of the social contract. Therefore, it is theoretically impossible to trigger the cheater-detection algorithm without presenting the cost/benefit relationship.

In light of this, another way must be found to induce the cheater-detection algorithm, or a similar Darwinian algorithm, without appealing to this cost/benefit structure. Consider, for a moment, the act of *cheating*, where cheating is defined in terms of taking a benefit without paying a cost. Cheating is merely one of several forms of *deception*. For instance, *lying* – the act of purposely imparting false information to another party – is another form of deception that does not (necessarily) involve a cost and a benefit. The *motivation* for a particular lie may be to cheat on a particular social contract, but the lie may be concocted without the need to refer to costs or benefits. Lying may serve other purposes as well, like avoiding blame or responsibility for an unfavourable outcome, or to hide one's intentions form an adversary. Lying is, on the other hand, a necessary part of cheating. When a character agrees to pay a cost on a social contract, but intends to do otherwise, they are, in effect, lying.

So, is it possible that the effect of facilitation on the Wason Selection task is really due to a more general mechanism of *deceit-detection*? Under such a hypothesis, cheater-detection would simply be a specific instance of deceit-detection. Unlike Cosmides, I make no claims about *how* such a mechanism would work at this stage – it is only suggested as a theoretical possibility that could also account for her results.

---

[1] Ignoring for a moment the experiments of Liberman & Klar (1996), and the criticisms by Cheng & Holyoak, (1989).

3.1 Experiment 1a:

The primary aim of this first experiment was to construct non-social contract scenarios that cued participants to check for instances of *deceit-detection*, and test them against social contract scenarios involving *cheater-detection*. The prediction (which served as a null-hypothesis) was that *there would be no significant difference between participants' performance on the Wason Selection Task under these two conditions*. This is because the deceit-detection hypothesis treats cheater-detection as simply a specific case of deceit-detection. On the other hand, Social Exchange theory argues that *the non-social contract deceit-detection tasks would induce significantly less facilitation than the social contract conditions* (which serves as an alternative hypothesis).

It was also decided that the non-social contract deceit-detection tasks (NSC deceit tasks, as an abbreviation) would be constructed to the prescriptions that Liberman & Klar suggest lead to good task understanding (and, therefore, high facilitation). This was used, for the most part, as a pretest to Experiment 2, (where one of the derivable predictions from Liberman & Klar's theory was tested). At first glance, this may appear to be a confounding influence, but this is not necessarily so. Liberman & Klar's theory and the deceit-detection hypothesis both predict that there should be no difference in participants' performance under social contract (SC cheating tasks) and NSC deceit tasks. Therefore, if the NSC deceit tasks were found to induce *lower* rates of correct responding, the deceit-detection hypothesis[1] would be falsified. This would, in turn, support Social Exchange theory. On the other hand, if there were no significant difference between SC cheating tasks and NSC deceit tasks, this would lend initial support to the deceit-detection hypothesis. Further tests would be needed to tease out the differences between the deceit-detection hypothesis, Social Exchange theory and Liberman & Klar's theory of Task Understanding.

Stimulus design:

Three groups of stimuli were created: SC cheating scenarios, NSC deceit scenarios and a thematic-abstract scenario. Three SC cheating scenarios and one thematic-abstract scenario were taken from the literature with a few minor alterations. The NSC scenarios were constructed from scratch. (*All the tasks are given in Appendix A2.1*).

---

[1] It might also be claimed that this would also invalidate Liberman & Klar's theory. But as it's inclusion in this part of the study is only a side issue, and for other reasons elaborated on in the discussion to this experiment, this is not necessarily so.

The thematic-abstract task was essentially the same as used by Cosmides (1989, p.192). Two cosmetic alterations were made to make this scenario more contemporary and coherent. Firstly, the task tells the participant that the ratings and codes are to be *entered into the school database*. Cosmides' tasks simply told her participants that the documents were to be "processed" without indicating what this meant. Secondly, a rational was presented which served to explain why the participant should believe some of the documents might have been incorrectly processed: *"(apparently, the previous secretary was often unreliable)"*. Neither of these alterations was predicted to have any effect on responding. This task was included to give a baseline for "abstract responding" and as a comparison to similar tasks in the literature.

The SC cheating scenarios were taken from a number of sources in the literature. The "*Drinking Age rule*" (Griggs & Cox, 1982) was included because of its historical significance and it's reliability. Throughout the literature it regularly produces rates of correct responding of 75% or more. It was adapted to be consistent with New Zealand drinking laws. In the traditional version of the task, the scene is set in a pub or bar. In many overseas countries people under the drinking age limit are allowed into pubs or bars, so long as they are not drinking and fulfil certain prerequisites, such as being accompanied by a parent or guardian. Hence, in these cultures it is perfectly legal to find under-ages in these establishments. In such cases, law enforcers need only check people who are underage and drinking alcohol (**P&~Q**). This is not the case in New Zealand. Technically speaking, under-ages are not allowed into pubs or bars at all. In this case, a law enforcer would have to check all people who are underage no-matter what they were drinking (**P, ~P &Q**). However, New Zealand law allows for people of all ages into licensed restaurants. People over the aged 18-20 are allowed to drink alcohol provided they are consuming a meal, and people over 20 are allowed to drink alcohol regardless of whether they are eating a meal or not. Therefore, these alterations were made to the "Drinking Age" rule to make it consistent with New Zealand law. Although this made the scenario a little more complicated than usual, the meaning of the tested rule remained intact. These alterations were not predicted to have a significant effect on task performance.

The next SC cheating scenario was the "*Overnight rule*" adapted from the text of Gigerenzer & Hug[1] (1992, p.142; Liberman & Klar, 1996, p.150). The scenario was constructed in accordance with the summary given by Gigerenzer and Hug, but because it was not copied there are almost certainly cosmetic differences between their original version[2] and the reconstruction used in this experiment. Again, these differences were not expected to effect participants' task performance.

The last of the SC cheating scenarios was an adaptation of the "*Grover High School problem*"[3] (Cosmides, 1989, p.268; Gigerenzer & Hug, 1992, p.142; Liberman & Klar, 1996, p.138). A number of changes were introduced to make the content relevant to New Zealanders.

Firstly, the "*women*" volunteers were replaced by "*parents*" to avoid any gender bias.

Secondly, the "*towns*" of Grover, Hanover and Belmont were replaced by "*suburbs*". This was to avoid the pragmatic improbability (at least, in New Zealand) of sending a child from one town to a state school[4] in another town. Most New Zealanders are familiar with the suburban zoning system used to allocate pupils to the closest state schools. To remain consistent with the switch to suburbs, the way in which the parents "paid the cost" in order to send their children to Grover High, was changed from "*through taxes*" to "*through city rates and community funding*".

Thirdly, the benefit for going to Grover high ("*an excellent record for getting students placed in good colleges*") was replaced by suggesting the school had "*excellent facilities and a superior academic record*". This was viewed as necessary because there is not so much variation in New Zealand with regards to the quality of universities (compared to the USA). The secondary school a pupil attends has no effect on their acceptance to a New Zealand university – only his or her academic success in nation wide exams determines entrance. Likewise, the problems associated with going to Hanover High school ("*poor teachers and decrepit facilities*") was replaced with "*decrepit facilities and a low academic record*", to remain consistent with the Grover High benefit.

Finally, a number of the passages were reordered. Any reordering was performed within the original paragraph. This served two functions. Firstly, in the first paragraph, the volunteers' task was introduced earlier to give the reader a better understanding of the rest of the passage. Secondly, the final paragraph (explaining the cards and the

---

[1] The original text is given in Appendix A1.7.
[2] Because Gigerenzer & Hug's participants were Austrian, the scenarios were probably not written in English.
[3] The original text is given in Appendix A1.5.
[4] This sort of thing is quite common in New Zealand for *private* schools, but the schools mentioned in this scenario are state schools run by the Board of Education.

task the participant has to achieve) was reordered to remain consistent with all the other scenarios in the experiment.

Again, none of these changes were expected to change the overall level of performance of participants. They should counter any loss of performance that might be expected, due to cultural differences, from the application of a United States designed task to New Zealanders.

Once the thematic-abstract and SC cheating scenarios were selected, the NSC deceit tasks were concocted. Previously in the literature, it has been usual, when comparing different types of tasks (say, social contracts and NSC permission rules) to use similar contexts and test rules of the same wording. It soon became apparent that it was too difficult to adapt SC cheating tasks to NSC deceit tasks. This is because the tested rule in a social contract task, by definition, contains the now familiar cost/benefit structure. Therefore, utilising the same rule for a NSC deceit task would be problematic. If there was no difference in performance between the SC cheating and the NSC deceit task, Social Exchange theorists could claim this is because *both* tasks elicit the cheater detection algorithm. Even if the costs and benefits were removed from the rule by redefining them in the text of the scenario as propositions without values, people might still incorrectly *interpret* the rule as having a cost/benefit relationship. With this in mind I decided to create completely new scenarios for the NSC deceit tasks, which could not be realistically interpreted as a SC cheating task.

It should be noted that nearly every word in these NSC deceit tasks was included to (a) cue the participant to the tasks of deceit detection, or (b) construct the task to the prescriptions of Liberman & Klar (1996).

The first of the NSC deceit scenarios was the "Murder Investigation" scenario. This was designed to cue the participant to the perspective of a police inspector who believes a suspect might be lying in order to get off a murder charge. The tested rule was "*If the suspect's finger-prints are on a glass, then that glass must have contained red wine*". There are no explicit cost or benefits associated with this rule, merely a relational statement linking two propositions. It is, though, a feasible way to test the suspects' alibi.

The second NSC deceit scenario was the "Cookie Jar Caper" scenario. This cued the participant to the role of a parent who suspects that their six-year-old son is trying to cover up an episode of illicit biscuit consumption by blaming the family dog. The rule was "*If any biscuits are missing from a given jar, then its top must be off*". Again, there

are no costs or benefits embedded in the rule, but the linked propositions act as a test of the son's story.

The last NSC deceit scenario was the "Missing Pharmaceuticals" scenario. This was subtly different from the previous NSC deceit tasks. The other NSC deceit scenarios had one main suspect whom the participant checks for instances of lying. The "Missing pharmaceuticals" scenario had no individual suspects but, rather, anybody from a group of potential suspects could be the culprit. The rule being tested was not from a verbal testament, but from the consequences of a set of actions. The scenario cued the participant to the position of production manager of a pharmaceutical company who suspects that one of his staff is stealing items from the factory floor. The rule was "*If a container is labelled 'Passed', then it must come out of the 'full weight' chute*". Again, the rule contained no explicit cost or benefits but, if it were falsified, the outcome would suggest that deceit had occurred.

Participants:

An a priori power analysis was performed to gauge the number of subjects needed to establish statistical reliability. A number of non-parametric tests were to be used, the least powerful of which was a chi-square. Hence the power analysis performed was for a chi-square. This would leave more than enough power for the other tests (which were to be within-subject design tests). From general surveillance of the literature, I expected to find a rate of correct responding to the SC cheating tasks of around 75%, and I estimated that a drop in performance to 60% was a forgivable allowance for the NSC deceit tasks. This gave a moderate effect size w=0.35, and with a power P=0.95,[1] this suggested that 107 participants were needed. For counterbalancing purposes, 108 participants was decided to be a good number to work with.

The 54 males and 54 females were recruited by poster advertisement. All the participants went into a sweepstake for NZ$50.

Recruitment took place at "*Student Job Search*". Student Job Search is located in the university's Student Association building, and is a government funded organisation dedicated to advertising jobs for students. Over the summer break, all students who were receiving an unemployment benefit had to check-in with Student Job Search every two weeks to "prove" that they were actively seeking summer employment. This provided a number of benefits in recruiting participants. Firstly, there were a high number of people who passed through every day. Secondly, many of them were interested in the financial incentive of the sweepstake on offer for participants. Thirdly,

[1] On reflection, this was probably too high. P=0.8-0.9 would have been sufficient.

the population base was broader than, say, an introductory psychology class, because it included students from other course backgrounds and other tertiary institutions.

Two criteria were set for peoples' inclusion in the experiment. Firstly, to avoid testing people who might already have experienced the Wason Selection Task, people were excluded if they had taken courses that may have taught the task as part of their syllabus. These included any course on cognition or any advanced philosophy course (second year or above). Persons who claimed they had seen the task before were also excluded. Secondly, to ensure that participants had sufficient understanding of the text, only people who spoke English as their first language, or had resided in New Zealand or another English speaking country for five years or more, were included.

Two participants (one male and one female) were removed from the study for taking too long (over 20mins) to complete the tasks. They were also seen to be repeatedly communicating with each other and not concentrating on the tasks. They were replaced by other participants.

Apparatus:

Each participant was given a booklet containing five pages. The first page was an information and instruction sheet (shown in Appendix A2.3). The second page was a consent form for the participant to sign and date. The third, fourth and fifth pages were tasks from the three conditions, one task per page. Participants were given one scenario from each condition. The orders of the task conditions (6 combinations) and the scenarios from each condition (9 combinations) were counterbalanced giving 54 separate combinations of task conditions and scenarios. It was also decided to test males' performance against females' performance, so 'sex' was counterbalanced for task conditions and scenarios. This doubled the number of combinations to 108.

The arrangement of the "cards" at the bottom of each copy of each scenario was randomised for one sex then an exact replica was made for the second sex. In other words, males and females received precisely the same scenarios, in the same orders, with the same arrangements of "cards" at the bottom of each scenario.

Procedure:

A desk was set up at Student Job Search and a number of posters were displayed around the area advertising the study and the associated sweepstake. Interested people approached the desk where the experimenter briefly explained the task and ascertained whether the volunteer was qualified to do the experiment. If the volunteer

passed the criteria, they wrote their name and contact details on a sheet of paper for the sweepstake and were issued with a task booklet.

The participant completed the task booklet in the Student Job Search room before handing it back to the experimenter. After all booklets were completed, and replacements achieved, the sweepstake was drawn and NZ$50 given to the lucky participant.

Results:

Most participants took between 5 and 10 minutes to complete the three tasks. Nobody took longer than 15 minutes (except the two participants who were excluded). All participants circled at least one answer to every question.

The results were analysed on two levels. Firstly, each participant's response was coded correct (**P&~Q**) or incorrect (any other response) for each scenario. Secondly, the actual pattern of responses (the combinations of **P, ~P, Q,** and/or **~Q**) was examined for each scenario. This serves as an analysis of the pattern of mistakes.

Table 3.1 shows the percentages of correct responses for each scenario, for sex of participant, and for each condition. The total number of correct responses for each category are shown in brackets.

Table 3.1 Percentage of correct responses[1] (P&~Q) for Task Conditions, Individual Scenarios and Sex of Participant.

Task Conditions	Scenario	Scenario Performance (N=36)	Sex		Task Condition Performance (N=108)
			Males (N=54)	Females (N=54)	
Thematic-abstract			15% (8)	4% (2)	9% (10)
SC cheating	Drinking Age Rule	64% (23)	65% (35)	50% (27)	57% (62)
	Overnight Rule	61% (22)			
	Grover High School Rule	47% (17)			
NSC deceit	Murder Investigation	44% (16)	41% (22)	33% (18)	37% (40)
	Cookie Jar Caper	41% (15)			
	Missing Pharmaceuticals	25% (9)			
		Means (M/F): N=162	40% (65)	29% (47)	

[1] All results are rounded to the nearest full percentage.

A Cochran's "Q" test for dependent samples[1] was used to test for significant differences between the task conditions (Siegel, 1956). The differences between all three tasks were found to be highly significant:

Thematic-abstract vs. SC cheating:	Q=52.00, df=1, p<0.001.
Thematic-abstract vs. NSC deceit:	Q=25.00, df=1, p<0.001.
SC cheating vs. NSC deceit:	Q=11.52, df=1, p<0.001.

Results were also analysed by sex-of-participant. Differences were tested for significance by two-tailed Chi-square. A significant difference was found when the performances of the males and females were compared across all three task conditions. However, this sex difference was only statistically reliable for the Thematic-abstract condition:

Thematic-abstract:	males 15%, females 4% (χ^2=3.97, df=1, p<0.05).
SC cheating:	males 65%, females 50% (χ^2=2.42, df=1, *not significant*).
NSC deceit:	males 41%, females 33% (χ^2=0.64, df=1, *not significant*).
All task conditions:	males 40%, females 29% (χ^2=4.42, df=1, p<0.05).

Performances for each of the SC cheating and NSC scenarios were examined. The lowest performing scenarios from each condition (the "Grover High School" and the "Missing Pharmaceuticals" scenarios) were tested against the other two scenarios within the same condition to see if they induced significantly lower performances.

"Grover High School" vs. "Drinking Age rule" and "Overnight rule"	(χ^2=2.29, df=1, not significant);
"Missing Pharmaceuticals" vs. "Murder Investigation" and "Cookie Jar Caper"	(χ^2=3.36, df=1, not significant[2]).

Next, the pattern of responses was examined for each scenario. Figure 3.1 shows the tallies of the responses given for the thematic-abstract scenario. Eleven of the fifteen[3] possible combinations of responses are represented here. Of particular note are the first four categories "**P&~Q**", "**P&Q**", "**P only**", and "**All four**". The "**P&~Q**" category is the correct response. As shown in the graph, this is rather low, indicating the poor performance of participants on this task (9%). Most of the errors were

---

[1] A McNemar's Chi-square test could also have been used, but Cochran's "Q" was easier given the tabulation of the data, and is a more powerful test.

[2] In fact, p=0.067. This value is included to show that the difference was approaching statistical significance. Indeed, if either the performance of the "Missing pharmaceuticals" scenario was *any* lower, or if the performance of one of the other NSC deceit scenarios was *any* higher, then the difference here would have reached significance. The problem here is one of inadequate power for this comparison.

[3] This number would be 16 if you included the option of *not* turning over *any* cards.

produced in the "**P&Q**" and "**P only**" categories, which are the common errors in the literature for pure abstract tasks. As reported in Chapter 1, the high "**P&Q**" category is indicative of the "matching bias" effect found in pure abstract tasks and thematic tasks that have very abstract content. This category dominates the response pattern for the thematic-abstract scenario recorded twice as many times (39%) as the nearest category "**P only**" (19%). The other errors are too low to be of any interest.

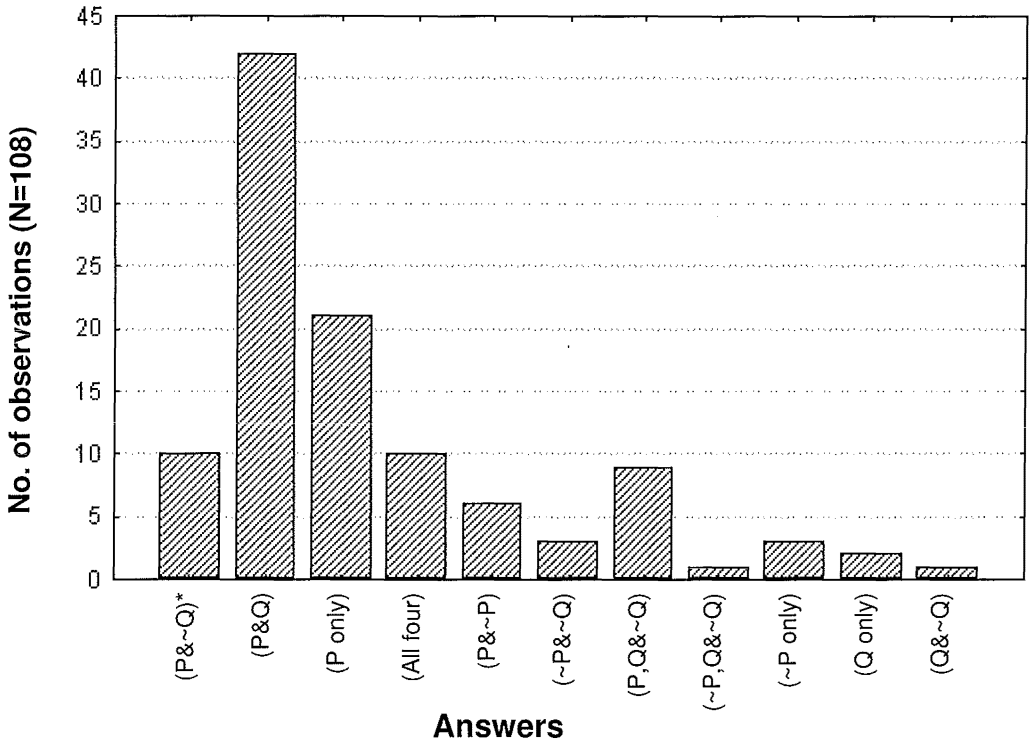


Figure 3.1: Graph showing the pattern of responses to the Thematic-abstract scenario in Experiment 1a. The correct answer is shown by *.

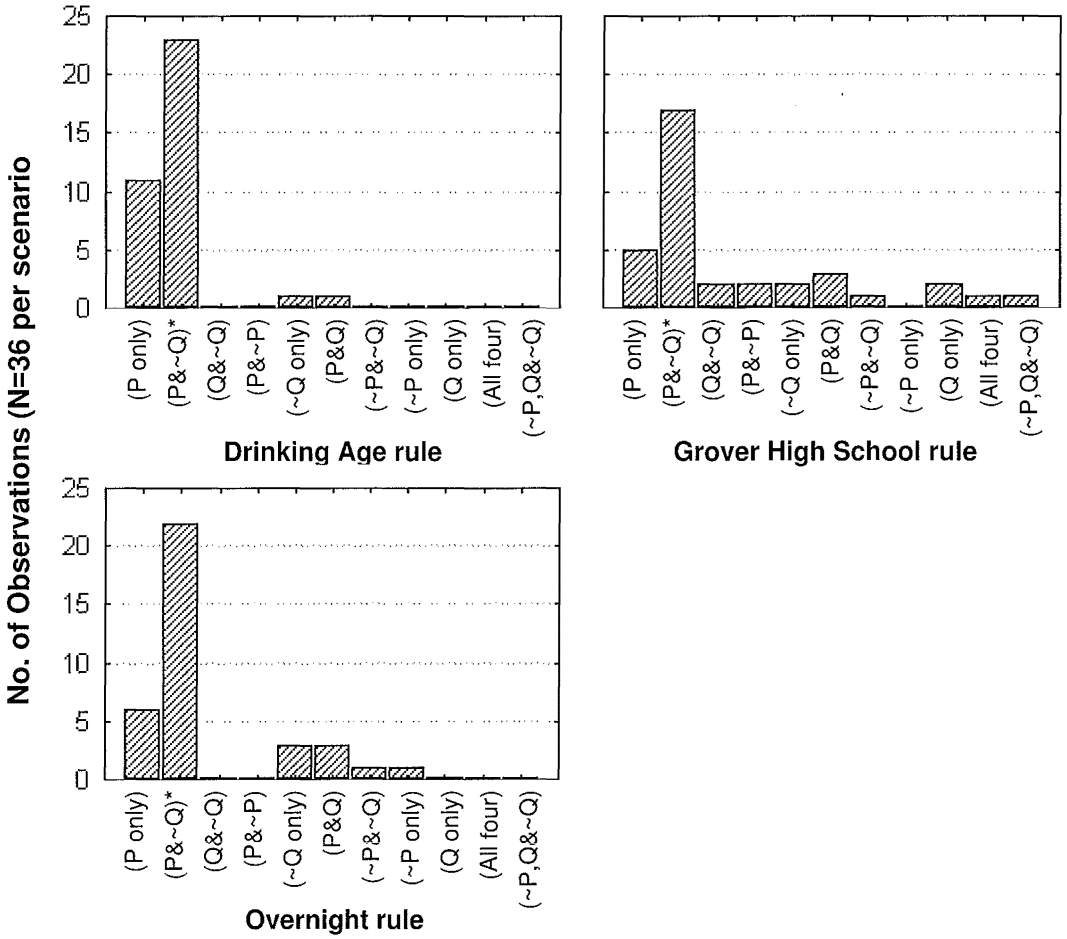


Figure 3.2: Graphs showing the pattern of responses to the three Social Contract cheating scenarios in Experiment 1a. The correct answer is shown by *.

Figure 3.2 shows the tallies of the responses given for each of the SC cheating scenarios. The most salient feature of these sets of graphs is the high number of the correct "**P&~Q**" responses, indicating the relatively good performance of the participants on the SC cheating tasks. Of the errors made, only the "**P only**" response is high enough to warrant mentioning. Again, this response is traditionally the most frequent error on SC cheating tasks, along with "**P&Q**" (Gigerenzer & Hug, 1992). Also of note is the specific error pattern of the "Grover High School" scenario. Although the level of correct responding is lower than the other two SC cheating scenarios, the corresponding increase in errors is spread across the error categories. This indicates there is no particular error bias associated with this scenario.

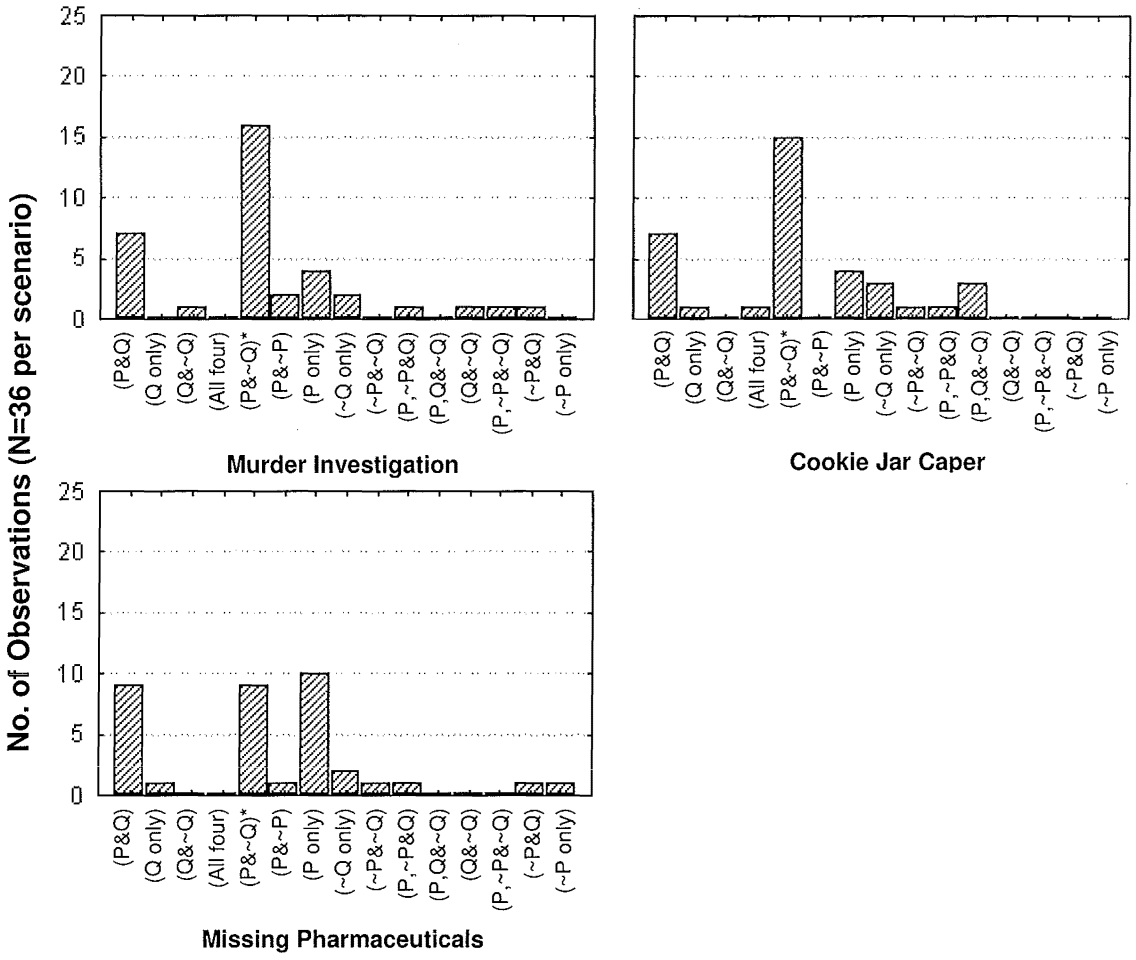
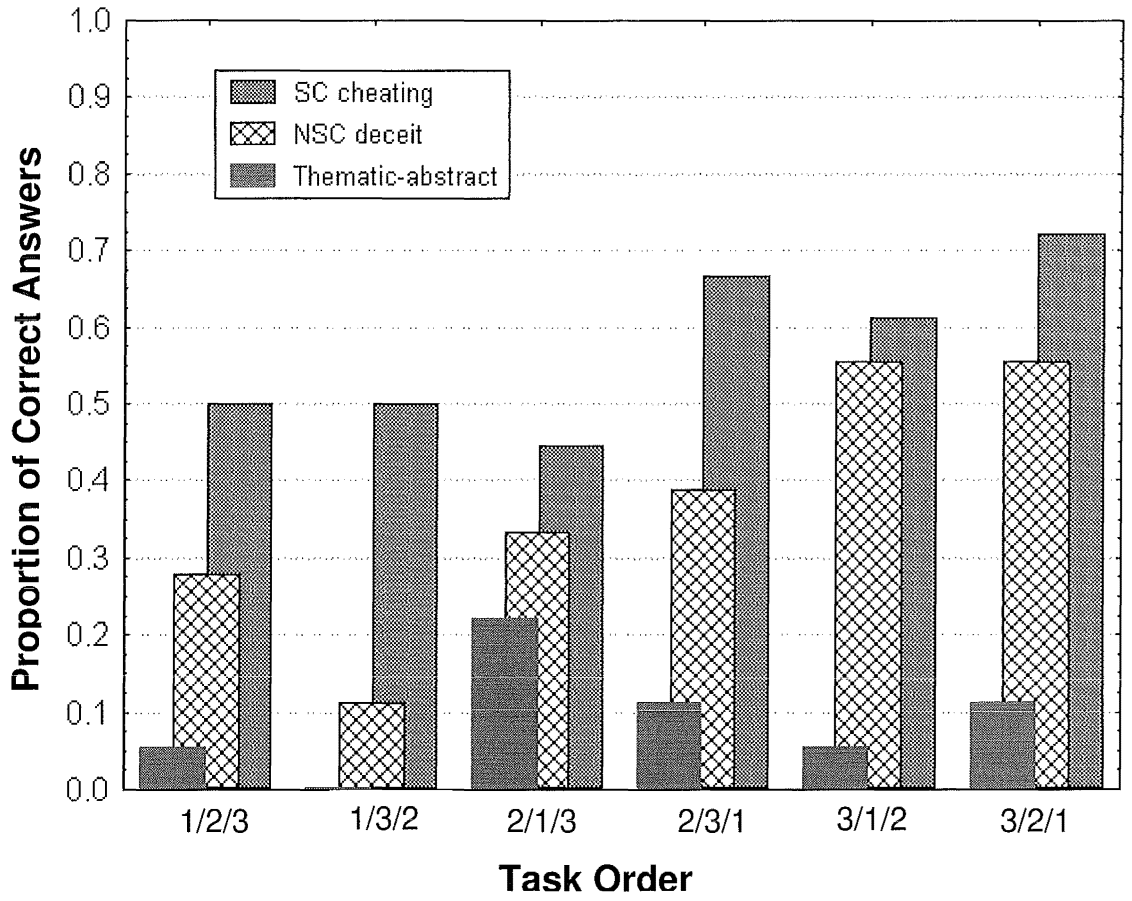


Figure 3.3: Graphs showing the pattern of responses to the three Non-Social-Contract deceit scenarios in Experiment 1a. The correct answer is shown by *.

Figure 3.3 shows the tallies of the responses given for each of the NSC deceit scenarios. The "Murder Investigation" and "Cookie Jar Caper" scenarios exhibit very similar response patterns. The correct "**P&~Q**" answer is the most popular response (44% & 42%); more than double the highest error response, "**P&Q**" (19%). All the other error categories slow fairly low frequencies, including the "**P only**" category (11%). This shows that although a large minority of people seemed to understand the task, a few people may have resorted to matching bias (indicated by the "**P&Q**" category). The "Missing Pharmaceuticals" scenario showed subtle differences from the other NSC deceit scenarios. In particular, a large minority of people chose the "**P only**" response (28%), more so than gave the correct "**P&~Q**" response (25%). The "**P&Q**" (25%) was also relatively high in comparison to the other error responses. This scenario is hard to interpret – it seems to show an error pattern closer to the thematic-abstract task than to

the other NSC deceit scenarios. This is supported by, or perhaps because of, the lower rate of correct responding.



(1=Thematic-abstract, 2=SC cheating, 3=NSC Deceit)

Figure 3.4: Graph showing the performances of the three task conditions for the six combinations of Task Order. The first number in each category label indicates the first task presented, the second number indicates the second task presented, and the third number indicates the third task presented.

Finally, an exploratory visual analysis of Task Order (the order in which each task condition was presented) was performed. Figure 3.4 shows the proportion of correct answers for the three task conditions (Thematic-abstract, SC cheating and NSC deceit) for the order each task condition was presented. The numbers in each category on the X-axis indicate the order of each task condition, with the left most number indicating the first task condition presented, the middle number indicating the second task condition presented, and the right most number indicating the last task condition presented. For example, the third category on the X-axis, "2/1/3" indicates that the first tasks presented were the SC cheating tasks ("2"), the second task presented was the

Thematic-abstract task ("1"), and the third tasks presented were the NSC deceit tasks ("3").

The graph shows that the relative performances of both the Thematic-abstract task and the SC cheating tasks vary little with respect to the task orders. The performances of the NSC deceit tasks, though, appear to fluctuate in accordance with *task order*. In particular, the NSC deceit tasks appear to elicit *lower performance* when they are *preceded* by the thematic-abstract task (as are the cases in the 3 left most categories), and elicit *higher performance* when they *precede* the thematic-abstract task (as are the cases in the 3 right most categories). This effect was robust enough to stand up to a test of statistical significance:

Mean percentage correct of the 3 NSC deceit tasks which *precede* the thematic-abstract task, (m_1=50%) vs.
Mean percentage correct of the 3 NSC deceit tasks which *are preceded by* the thematic-abstract task, (m_2=24%);
χ^2=7.78, N=108, df=1, p<0.01.

This result should be accepted with a certain amount of caution, however, as it is the outcome of a post-hoc analysis.

Discussion:

The main hypothesis of Experiment 1a was that there should be no significant difference between the SC cheating tasks and the NSC deceit tasks. This proved to be false. Moreover, the results support Social Exchange theory's prediction that the SC cheating tasks should perform better than both the NSC deceit and the Thematic-abstract tasks. It seems evident, then, that *the possibility of thematic facilitation due to simple deceit detection must be rejected.*

However, there is also weak evidence for Liberman & Klar's theory of task understanding. Although performance on the NSC deceit tasks was lower than the SC cheating tasks, performance on the NSC deceit was also significantly better than on the Thematic-abstract task. This second finding was predicted by Liberman & Klar's theory, but not by Social Exchange theory.

The situation is further complicated when the performance of each individual scenario is taken into account. Although the "Drinking Age rule" and the "Overnight rule" elicit reasonably good performance, the other "SC cheating" scenario, the "Grover High School" scenario elicited considerably poorer performance (about 15% lower). Likewise, the "Missing Pharmaceuticals" scenario elicited a lower success rate (by about 18%) than the other NSC deceit scenarios. The difference in success between the means of the SC cheating and the NSC deceit conditions was about 20% This suggests that the within-condition performance was nearly as varied as the between-

condition performance. Indeed, the "Grover High School" scenario did not appear to elicit any better performance than the "Murder Investigation" or "Cookie Jar Caper" scenarios. Such within-condition variance is better explained by *task understanding* than by *Social Exchange theory.*

The theory of task understanding suggests that certain elements of each task will contribute to participants' understanding of and, hence, their performance on the task. This suggests that each scenario should be examined *individually* with respect to the elements that enhance or inhibit task understanding and, hence, performance. Therefore, it predicts that there can be a large amount of within-condition variation. Analysis of Liberman & Klar's (1996) experiments confirms this: in their first experiment, original SC cheating scenarios varied in success from 52-81%, while the unconfounded SC no-cheating scenarios varied from 58-85%. Social Exchange theory, on the other hand, suggests that the only relevant factor in determining task performance is the presence or absence of a social contract which triggers the cheater-detection algorithm[1]. Therefore, Social Exchange theory predicts very little within-condition variance.

One of the most perplexing features of the results in Table 3.1 was the relatively low rates of correct responding across all three task conditions. Normally, social contract tasks elicit rates of between 70-80%[2], and thematic-abstract tasks usually elicit rates of 15-20% when included with facilitating tasks. Hence, the rates of 57% for the SC cheating and 9% for the thematic-abstract task were lower than expected.

There are three conceivable reasons for the low rate of overall correct responding. Firstly, there are a number of possible population differences between the sample used here and other samples from the literature. For instance, the general level of intelligence may be lower in my sample than in previous samples. As mentioned in chapter 1, IQ is a positive predictor of task performance, and it is probable that many of the samples in the literature were from populations of higher academic ability[3] than found at Student Jobs Search[4]. Secondly, the test environment was not ideal. There was a significant amount of heavy construction going on in the Students Association Building over the summer break. Hence, there was often an unanticipated level of

---

[1] It has been argued that the cheater-detection algorithm might be triggered more easily by some social contract than others. This *would* explain the large within-condition variance. However, it also begs the question "*Why* are some social contracts better at triggering the cheater-detection algorithm than others?" Some thing *other* than cheater-detection must be causing this. Task understanding, perhaps?

[2] However, success rates as low as 52% have been cited for single SC cheating scenarios (see Liberman & Klar, 1996).

[3] For instance, Cosmides' samples were volunteers from Harvard and Stanford universities.

[4] This may be particularly true in light of the fact that a number of the participants were from polytechnics and training colleges.

background noise occurring while participants completed the tasks. This may have disturbed their concentration. Thirdly, there may have been deficiencies in the task construction. As mentioned in the stimulus design section, the SC cheating and Thematic-abstract tasks were either altered in some (supposedly cosmetic) way or reconstructed from a summary. It is possible that differences from the original scenarios were detrimental to task performance in unanticipated ways.

A significant sex effect was also evident in Experiment 1a. Males were found to do better on all three task conditions, although only the thematic-abstract task showed a statistically significant difference. However, when the results from all task conditions were pooled, the overall differences proved a significant 10% difference in performance. In terms of actual numbers correct, males were 38% higher (65 vs. 47) than females. A sex based performance difference such as this has not been reported in the literature before. David Hardman (personal communication, 17 January, 1997) has suggested that this may simply be because nobody has *looked* at sex differences before[1].

The cause of these differences is a mystery. There is no evidence that the males from the sample are any more intelligent than the females. It is also unlikely that the scenarios themselves were gender biased. Particular attention was paid in the construction of the scenarios to avoid stereotypes or sexist language. The sex of each character, who's perspective the participants were being cued to, was purposely left ambiguous in order that they could be interpreted as either male or female.

One possible reason for the observed sex disparity is cultural gender bias. Males in New Zealand and other western cultures have been stereotypically biased to think they should be "logical" and "mathematically orientated". The cultural stereotype for females is that they should be "intuitive" and "artistically inclined". Therefore, males have been prompted to be rational thinkers and have been rewarded for their logical achievements. Females, on the other hand, have possibility had less encouragement and reward for similar endeavours. Hence, males may have developed better reasoning skills via different learning conditions. They may, for instance, tend to examine the structure of the tested rule, whereas females may tend to rely on pragmatic or heuristic approaches. University facilities such as engineering, mathematics, computing and the physical sciences currently attract a significantly higher percentage of males than females. These courses tend to teach and require relatively strong logical reasoning skills, more so, perhaps, than arts and other courses.

---

[1] It seems difficult to believe that such an obvious line of research has not been investigated before. This makes one wonder whether such examinations *have* been performed, but not *reported* because of the current climate of political correctness.

Therefore, a kind of population gender bias may have occurred because males are more likely to come from these academic backgrounds.

The pattern of errors for the NSC deceit tasks also warrants discussion. As shown in figure 3.3, the "Missing Pharmaceuticals" scenario shows a pattern of results which is similar to the Thematic-abstract task as well as the other NSC deceit tasks. In particular, the high level of the "**P only**" answer differentiates this task from the other NSC deceit scenarios (in addition to the lower rate of correct responding). Re-examining the scenario, I suggest two possible reasons for the poor performance.

Firstly, the participant may not interpret the rule as *determinative* but, rather, *probabilistic*. It is mentioned near the start of the text that the manager of the pharmaceutical company expects *some* products to be underweight (about 1 in 10,000) and escape detection. Hence, technically speaking, the rule *is* probabilistic, but as the chance of detection failure is so low, the rule could be thought of as totally deterministic for practical reasons. Nevertheless, participants may still take the probabilistic nature of the rule into account when considering whether to select the "**~Q**" card. If this is right, then the participants are *correct* in not choosing the "**~Q**" card, because an "underweight container" (**~Q**) which had a "passed label" (**P**) could have been mis-weighed in the original pass through the sorting machine. The problem with this hypothesis is that, logically speaking, selecting the "**P**" card should not falsify the rule either, for the same reason as the "**~Q**" card. In fact, given the probabilistic hypothesis, there is no (guaranteed) way of falsity the rule at all!

The second possibility for the poor task performance is that participants do not fully understand the nature of the rule because of the shear length of the scenario. It contains about 40% more words than Cosmides' version of the "Grover High School" scenario. It also makes for quite dense reading: each sentence and proposition is necessary for full task understanding. Hardman (1996) has commented on the considerable load the pure-abstract task places on working memory. It is probable that this scenario contains too many propositions for participants to process properly to obtain a good understanding of the rule being tested. As a corollary observation; participants often complained that they found this task confusing and "too much to handle".

The results of the post-hoc analysis of the Task Order are also intriguing. The order in which the conditions were presented had a significant influence on the performance of the NSC deceit scenarios, but not on the thematic-abstract or SC cheating tasks. In particular, performance on the NSC deceit seemed to suffer if it was preceded by the thematic-abstract task. This might be explained by appealing to the phenomenon of reasoning-by-analogy. If we assume that poor performance on the thematic-abstract task is caused by poor task understanding, then a participant may apply an erroneous strategy to answer the task. Then, when the participant is faced with the NSC deceit task, the heightened availability of the erroneous strategy means the participant is more likely to utilise the strategy to answer the problem.

This raises questions about the other task conditions. If performance on the NSC deceit task is contingent on the preceding task, then why is this not the case with the Thematic-abstract or SC cheating conditions? It is well documented in the literature that thematic-abstract and social contract tasks appear to be remarkably resistant to reasoning-by-analogy (Evans et al., 1993). This poses a problem for the theory of Task Understanding, because it predicts that how one understands one task should influence the understanding of another task, if only as a minor tendency. Another problem is that the only the thematic-abstract tasks appears to influence the NSC deceit tasks; there is no particular increase in performance when the NSC deceit scenarios are preceded by the SC cheating scenarios.

In light of the apparent inconsistencies, the finding that the performance of the NSC deceit tasks is contingent on the presence or absence of a preceding thematic-abstract task must be considered with an element of scepticism. The phenomenon requires further investigation; this will be examined in Experiment 1b.

3.2 Experiment 1b:

One of the auxiliary findings of Experiment 1a was that the task order had an apparent effect on the performance of the NSC deceit tasks. In particular, the NSC deceit scenarios elicited lower performance when they were preceded by the thematic-abstract task.

Because the NSC deceit scenarios were designed using the prescriptions Liberman & Klar suggest should lead to good task understanding, they should, in effect, be equivalent to Liberman & Klar's *unconfounded no-cheating* scenarios. Liberman & Klar found that these scenarios performed as well as SC cheating scenarios. Yet performance on the NSC cheating scenarios in Experiment 1a was significantly lower than on the SC cheating scenarios. There is, however, a discrepancy between the methodology of Experiment 1a and Liberman & Klar's experiment. Liberman & Klar's experiment did not include any kind of abstract task. If the finding that a preceding thematic-abstract task reduces correct responding on a NSC deceit task holds, then we should expect performance to improve when the thematic-abstract task is omitted.

Therefore, the aim of Experiment 1b is to partially retest the NSC deceit and SC cheating conditions without the thematic-abstract task. The prediction is that performance on the NSC deceit tasks should improve, and that there should be no significant difference between the NSC deceit task condition and the SC cheating condition.

Scenarios:

It was decided that only the two best performing scenarios from each condition in Experiment 1a should be used. This was because of the problems associated with the "Missing Pharmaceuticals" scenario, namely the scenario length and possible probabilistic interpretation of the rule (*see the discussion section of Experiment 1a*). As the lowest performing NSC deceit scenario was discarded, it was decided that the lowest performing SC cheating scenario should also be discarded to avoid biasing the task conditions. Therefore, the NSC deceit scenarios were the "Murder Investigation" and the "Cookie Jar Caper", while the SC cheating scenarios were the "Drinking Age rule" and the "Overnight rule".

Participants:

Since the prediction is that there should be *no* significant difference between the NSC deceit tasks and the SC cheating conditions, enough participants are needed to

ensure there is enough statistical power to find significant results. If there is no significant result, the possibility that there was insufficient power can be excluded. Unfortunately, it appears that no equation has been devised for a power-analysis of non-parametric dependent measure analyses such as the Cochran's "Q" test that will be used (Cohen, 1988). Therefore, it was decided to reanalyse the results from Experiment 1a, removing the "Missing Pharmaceuticals" and "Grover High School" results. This left 48 participants; the percentage of correct responses for the NSC deceit task condition was 42%, the percentage of correct responses for the SC cheating task condition was 65%. The difference was significant ($Q=5.76$, $df=1$, $p<0.05$). Hence, N=48 participants was considered adequate for Experiment 1b.

The participants were 48 volunteers from Student Job Search. These comprised 24 males and 24 females. The same criteria for exclusion from Experiment 1a were applied, with the addition that people who participated in Experiment 1a were also excluded. All participants went into sweepstake for NZ$50.

Apparatus:

Each subject received a four-page booklet containing an information sheet (*see Appendix A2.3*), a consent form and one scenario from each of the task conditions. The presentation of the scenarios was counterbalanced for task condition order. The order of the "cards" in each scenario was randomised.

Procedure:

The procedure was identical to that of Experiment 1a.

Results:

All participants completed the tasks within 10 minutes. No replacements were necessary.

Table 3.2 gives the percentage of correct responses for all four scenarios, for sex of participant, and for each condition. The total numbers of correct responses for each category are shown in brackets.

Table 3.2: Percentage of correct responses[1] (P&~Q) for Task Conditions, Individual Scenarios and Sex of Participant.

Task Conditions	Scenario	Scenario Performance (N=24)	Sex		Task Condition Performance (N=48)
			Males (N=24)	Females (N=24)	
NSC deceit	Murder Investigation	67% (16)	50% (12)	63% (15)	56% (27)
	Cookie Jar Caper	46% (11)			
SC cheating	Drinking Age Rule	67% (16)	71% (17)	71% (17)	71% (34)
	Overnight Rule	75% (18)			
	Means (M/F): N=48		60% (29)	67% (32)	

The SC cheating condition (71%) elicited better performance than the NSC deceit condition (56%), but this difference was not statistically reliable (Q=2.58, df=1, *not significant*). This lends putative support to the hypothesis that removing the thematic-abstract task increases NSC deceit performance to a level where it becomes indistinguishable from the SC cheating condition. However, when the performance of the NSC deceit tasks in Experiment 1b are compared to those from Experiment 1a[2] (42%), the difference is *also* not reliable (χ^2=2.04, df=1, *not significant*).

The comparison of the performances on each scenario proves interesting. Although the "Cookie Jar Caper" only elicited a 46% success rate, the other NSC deceit task, the "Murder Investigation" elicited a very creditable 67%. This was the same level of success as the SC cheating "Drinking Age rule". However, the second SC cheating scenario, the "Overnight rule" elicited even higher performance, at 75%.

The sex effect found in Experiment 1a was not found in this experiment. Although females (63%) out performed males (50%) on the NSC deceit tasks, this difference was not reliable (χ^2=0.76, df=1, *not significant*).

[1] All results are rounded to the nearest full percentage.
[2] This result only takes into account the instances of the "Murder Investigation" and "Cookie Jar Caper" scenarios that were *not* presented with the "Grover High School" scenario.

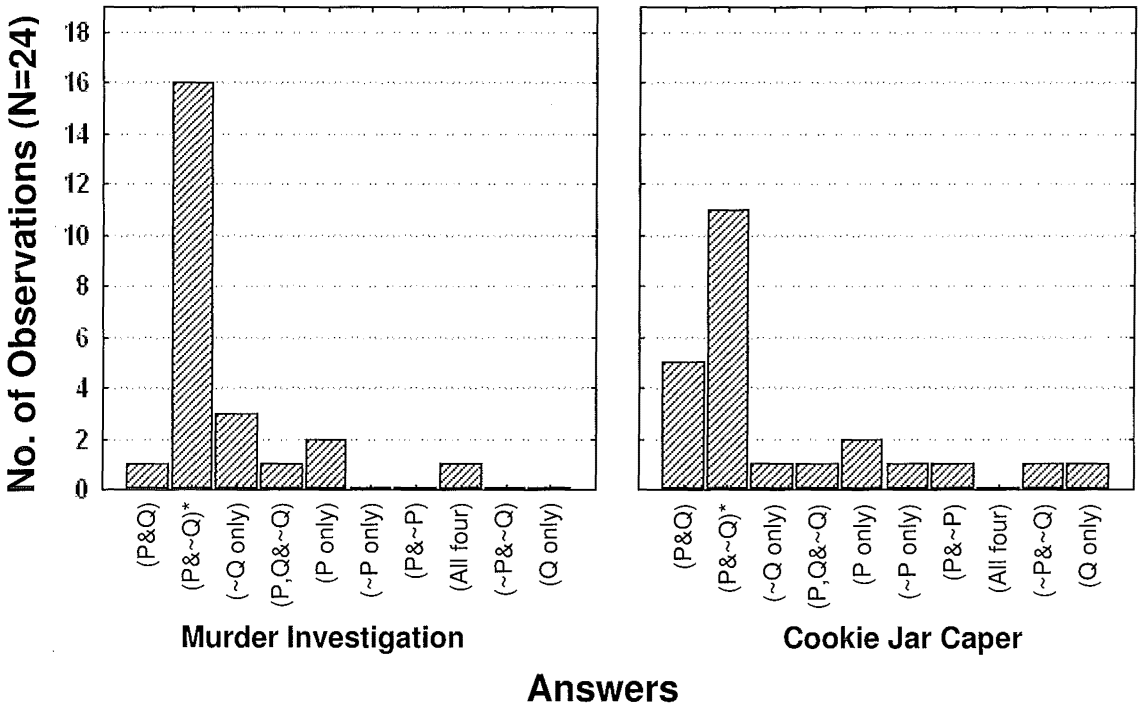


Figure 3.4: Graphs showing the pattern of responses to the Non-Social-Contract deceit scenarios in Experiment 1b. The correct answer is indicated by *.

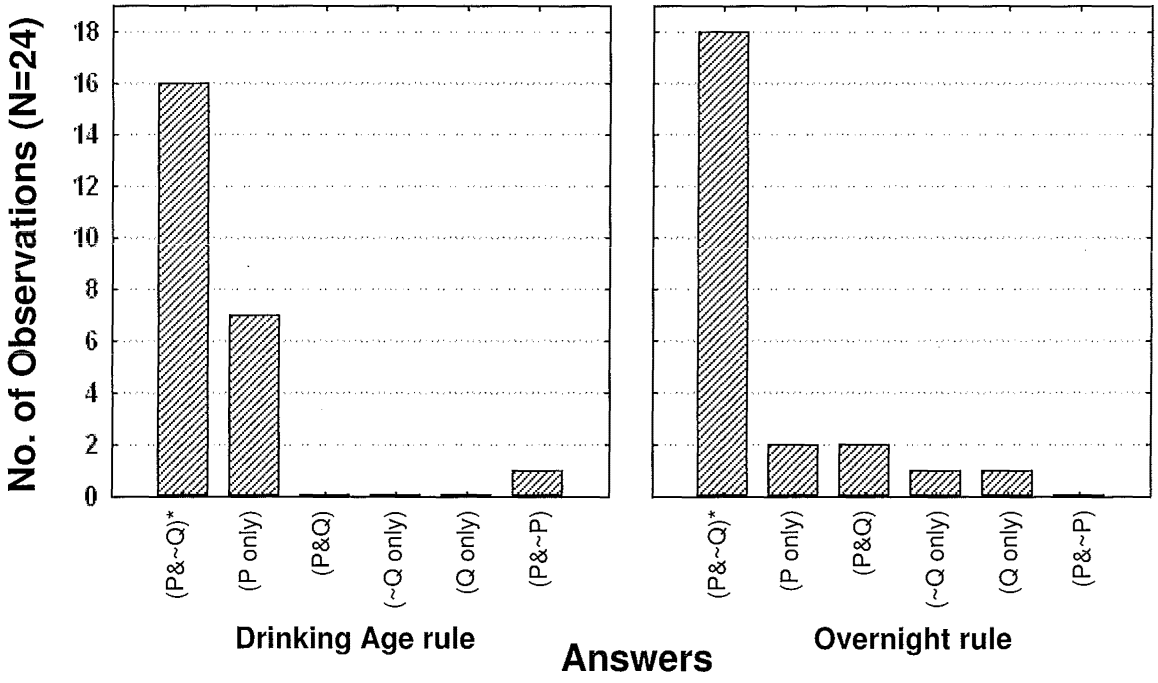


Figure 3.5: Graphs showing the pattern of responses to the Social Contract cheating scenarios in Experiment 1b. The correct answer is indicated by *.

Another analysis of errors was performed for all four scenarios. Figure 3.4 shows the pattern of responses for each of the NSC deceit tasks, and Figure 3.5 shows the pattern of responses for each of the SC cheating tasks. None of the graphs in these figures showed error patterns that are notably different to previous results. As in Experiment 1a, the "**P&Q**" and "**P only**" errors tended to be the most frequent, and the NSC deceit scenarios showed a wider variation than the SC cheating scenarios in the categories of errors.

Discussion:

Although the NSC cheating condition in Experiment 1b was not significantly lower than the SC cheating condition, it was also not significantly higher than the NSC cheating condition in Experiment 1a. Therefore, the hypothesis that removing the thematic-abstract task will improve performance on the NSC deceit tasks, while not rejected, cannot be supported by the data presented. The problem is one of statistical power.

A post-hoc analysis of the χ^2 between the NSC deceit conditions of Experiments 1a and 1b reveals that the power is slightly below P=0.8 (P=0.79; α=0.05, w=0.28, N=96). In fact, if only another 3 participants from Experiment 1b had got the NSC deceit task correct, then this result would have reached significance. It is apparent, then, that more testing is needed. Experiment 1b should be repeated, but with more participants.

3.3 General Discussion of Experiments 1 a & b:

The primary motive for the experiments in this chapter was to test whether the high facilitation found on social contract versions of the Wason Selection task could be attributable to general deceit detection, rather than specific detection of cheating. Experiment 1a appeared to rule this possibility out. Three NSC deceit tasks, which required no direct cheater detection, elicited significantly lower overall performance than three SC cheating tasks. There were problems with these results however.

Firstly, the overall level of responding was somewhat lower than expected. The cause of these lower than expected findings is uncertain, but is most probably due to differences in the characteristics of the sampled population compared to the populations normally used in the literature.

Secondly, although the performance on the NSC deceit tasks in Experiment 1a was lower than on the SC cheating tasks, it was also significantly higher than on the thematic-abstract tasks. This is difficult for Social Exchange theory to explain; it treats all non-social-contract tasks as equivalent. Ironically, the deceit-detection hypothesis cannot explain this result for much the same reason. It suggests that all deceit-detection tasks should elicit roughly the same levels of performance, including social contract cheating tasks. In this case, social contract tasks out-performed the NSC deceit tasks, so the hypothesis must be rejected.

The secondary motive for the experiments in this chapter was to serve as a pretest for Experiment 2. Experiment 2 is intended test one of the predictions of Liberman and Klar's (1996) Theory of Task Understanding. This theory suggests that each task should be considered on its own merit with respect to how well the participants understand certain features of the task. This explains the high within-condition variability compared to the between condition variability found in Experiments 1a and 1b. The SC cheating tasks which elicited the lowest performances failed to elicit better performance than the best of the NSC deceit tasks.

However, the NSC deceit tasks were designed using the prescriptions that Liberman & Klar suggest lead to good task understanding. Hence, they should, in theory, elicit equivalent performance to, if not better than, the SC cheating tasks. They did not. However, a post-hoc review of the NSC deceit tasks reveals that there were problems with these tasks that may have hindered task understanding.

The NSC deceit tasks appear to be longer, on average, than the SC cheating tasks. In particular, the "Missing Pharmaceuticals" scenario, which elicited the lowest

performance, was about 40% longer than the longest SC cheating task. It is probable that too much information was presented in this task for participants to successfully manage (particularly when the limited capacity of working memory is considered). This is not true of the other NSC deceit tasks, however. Although the "Murder Investigation" and "Cookie Jar Caper" scenarios are 50-100% longer than the "Drinking Age rule", they are roughly the same length as the "Overnight rule" and the "Grover High School" scenario.

Another difference between the SC cheating and NSC deceit conditions was the selection of the tasks. The NSC deceit tasks were ones that were designed specifically for this experiment and, therefore, had no previous history in the literature. The SC cheating tasks, on the other hand, were taken from the literature, and had a well-established history of success. It is possible, therefore, that the selection of these tasks biased the SC cheating condition in favour of success. It is quite possible, for instance, that other SC cheating conditions that have been used by previous researchers, but have not been reported in the literature because they performed lower than expected. This discrepancy will be rectified in Experiment 2 by devising new social contract scenarios to go with the new non-social-contract scenarios.

Experiment 1a also produced a significant sex effect. Generally, males seemed to perform better than females. This results was not stable, however; as this difference disappeared in Experiment 1b. Sex effects have not been recorded in the literature before, but this may simply be due to the fact that no one has *looked* for them.

This sex difference is hard to explain. None of the scenarios appear to be gender biased, and it is assumed that the males and females from the sample population were equivalent with respects to intelligence and educational level. Yet, it could be that the males in the sample tended to have developed better logical reasoning skills. Certain university facilities such as engineering, mathematics and the physical sciences still attract a much higher percentage of males than females. These courses tend to teach or require relatively strong logical reasoning skills, more so, perhaps, than arts and other courses. Hence, the sample population may be biased by the fact that males are more likely to come from these academic backgrounds.

A more likely explanation is that this finding is simply a statistical aberration. The fact that this difference was not evident in the Experiment 1b, even though the stimuli and test conditions were the same as in Experiment 1a, suggests that the result is not reliable. Only further testing will enlighten this possibility.

4 Experiment Two: Task Understanding vs. Social Contracts

4.0 Introduction:

Liberman & Klar's (1996) theory of Task Understanding suggests that performance on the Wason Selection Task is primarily due to participants' understanding of three fundamental aspects of the task (*see section 2.5*).

(1)	***The clarity of the rule in terms of determination and direction.***
(2)	***The nature of the alternative to the tested rule and the falsifying instance it entails.***
(3)	***The perceived relevance of looking for violations of the rule.***

A derivable prediction from this is that *all tasks, irrespective of context or content, could produce similar high levels of performance, if the three fundamental aspects are sufficiently understood.* In theory, even "*pure-abstract*" tasks (tasks presented without any thematic content) should elicit high levels of performance if they could be presented in a way that ensured that the three fundamental aspects were comprehended. This, in essence, is a testable hypothesis.

Normally, the three fundamental aspects are imbedded implicitly in the form of an accompanying scenario. Unfortunately, the very presence of a scenario turns the task into a thematic task. Yet, it is not obvious how this can be worked around. One cannot simply present a "cook-book" recipe for understanding the task, as this tantamount to giving the participant the answer. It is apparent, then, that thematic tasks must be used.

Thematic-abstract tasks, like the one used in Experiment 1a&b, are able to embed the three fundamental aspects within an accompanying scenario. Yet they still retain an element of abstractness by virtue of the fact that they do not conform to any other proposed task categories (such as "social contracts" or "permission and obligation rules").

Another method to improve task understanding might be to include a *visual* mode of presentation. The representation of a task in the form of pictures or diagrams may help comprehension in three ways.

Firstly, it may help consolidate the important aspects of the text. Often large amounts of verbally presented material are hard to maintain in working memory. A proposition read earlier in the text may be forgotten before the participant must call on it to evaluate the problem. An additional picture or diagram may provide a structural representation which provides the participant with an easy referent. The participant would need only to *look* at the picture to access the information, rather than having to reread the passage.

Secondly, it provides a means of presenting additional information that may be hard to verbalise. Occasionally, a concept is much easier to present in pictorial form than verbal form. In verbal form, such concepts may need long-winded explanations, or require complicated language which impedes understanding.

Thirdly, it may improve the clarity of the text by focusing the intent of the text and reducing ambiguity. Verbal language is, by nature, relatively ambiguous. Many words have double meanings, and many phrases will leave the reader with more than one possible interpretation. The inclusion of a pictorial representation may help clarify these instances.

4.1 Experiment 2:

There were two aims for Experiment 2. The first aim was to construct thematic-abstract tasks which conform to the three fundamental aspects of task understanding suggested by Liberman & Klar (1996), and to test their performance against similarly constructed social contract cheating-detection tasks. The hypothesis is that there should be no significant difference in the performances of the two task conditions. Social Exchange theory suggests that the social contract condition should out-perform the thematic-abstract condition. This is essentially a partial replication of Liberman and Klar's (1996) first experiment. The major difference is that they used the *same* tested rules in each condition, while manipulating the content of the scenarios to manufacture social contract and non-social contract versions of each rule. This led to scenarios that, while testing the same wording of the rule, referred to quite different versions that were only superficially related. Hence, it was decided here that different rules would be used in the different task conditions.

The second aim was to construct diagrams which demonstrated the structural relationship of the components in the tested rules for the thematic-abstract and SC cheating tasks, and test whether they have a significant effect on task performance. Half the scenarios would be presented with the text and a diagram, while the other half would be presented with the text only. The hypothesis was that the presence of the

diagram would significantly improve selection task performance for both thematic-abstract and SC cheating tasks.

Stimulus design:

It was decided to use *different* tasks to those used in Experiment 1. There were two reasons for this. Firstly, the NSC deceit tasks (which are effectively thematic-abstract tasks if one denies the validity of the deceit-detection hypothesis) were restricted by the initial need to produce tasks that conformed to deceit detection. This was a limiting factor, which at times conflicted with the incorporation of the three fundamental aspects of task understanding suggested by Liberman & Klar. Secondly, the SC cheating tasks were selected from the literature as tasks that performed well. Therefore, these tasks may bias the social contract condition because they are *known* to perform well. It is quite possible that there are social contract tasks that have performed poorly, but have not been published in the literature because they do not exhibit the performance that the experimenters expected. With newly created thematic-abstract tasks, there are no previous performance histories for comparison. By creating both types of scenarios from scratch, these possibilities are eliminated. In addition, any possible biases caused by the inability of the experimenter to create clear, concise scenarios[1], is also eliminated.

Three SC cheating tasks (the "*Abstract Social Contract*", the "*Product Promotion*" and the "*Security Guard*" scenarios) and three thematic-abstract tasks (the "*Engineering Diagram*", the "*Automated Rail Controller*" and the "*Smithton Swamp*" scenarios) were created (see Appendix A2.2). A picture or diagram was created for each scenario. Unlike the NSC deceit tasks of Experiment 1, the thematic-abstract tasks were not restricted to the detection of deceit.

The "*Abstract Social Contract*" scenario is essentially a 'bare-bones' social contract with all the contextual information removed. It uses abstract costs ("*tasks X and Y*") and benefits ("*Resources A and B*") without explaining what they are or do. But the relative magnitudes of the costs and benefits, and the structural contingency of the cost/benefit relationship are presented to the participant. Such a social contract scenario has not been presented before in the literature, although it is similar to Cheng & Holyoak's (1985) "abstract permission rule" (*see section 1.5*). There are distinct differences, though. As well as being a social contract rather than a permission rule, it also uses *implicit* negatives ("*task Y*" and "*resource B*") rather than *explicit* negatives ("*not task X*"

---

[1] For instance, one criticism that could be levelled at the NSC deceit scenarios in Experiment 1 was that they were too long.

and "*not resource B*"). Cheng & Holyoak's abstract permission rule was criticised by Jackson & Griggs (1990) for using explicit negatives. They found that when implicit negatives replaced the explicit negatives, performance on the task was low. This provides another testable hypothesis. If performance on the abstract social contract task is low (<20%), then this conforms to the findings of Jackson & Griggs[1]. Social Exchange theory, on the other hand, predicts that performance should be high (>50%), as activation of the cheater-detection algorithm should be context independent. The abstract content of the social contract alone should be enough to trigger the algorithm.

The diagram accompanying the "Abstract Social Contract" scenario shows the relationships between the tasks to be performed and the resources that *may* be taken. It does not, however, *explicitly* point out that "Resource A" *may not* be taken after performing "Task Y", as this might be viewed as giving the participant the answer[2].

The "*Product Promotion*" scenario is a standard thematic social contract problem. The participant is cued to the perspective of a shopkeeper who is worried that some customers may try to cheat on a product promotion rule. The benefit, in this case, is a free chocolate bar, while the cost is having to pay for a particular brand of soft drink. The accompanying picture depicts the store scene with display stands showing the particular soft drink brand and one of its competitors, and the chocolate bar. The cost/benefit contingency is spelt out on the advertisement for the soft drink, with an arrow pointing to the benefit.

The "Security Guard" scenario is another standard thematic social contract. The participant is cued to the perspective of a security guard working in a government building who is charged with ensuring that only qualified people are allowed into certain restricted areas. The benefit, in this case, is being allowed into a restricted vault (which contains valuable documents and information that other companies and governments would pay considerable money for), and the prerequisite (cost) is the possession for a high-ranking security pass (which must be earned through a long-time record of trustworthiness). The accompanying diagram shows the internal layout of the building, including the various security areas and the security pass code needed to be allowed into each one.

The "*Engineering Diagram*" scenario is the first of the thematic-abstract tasks designed to the prescriptions laid out by the theory of Task Understanding. The

---

[1] Jackson & Griggs' experiment found a mean of only 8% success across four versions of the tasks using implicit negatives.

[2] This is true of all the pictures and diagrams in the scenarios. Negative relationships were never made explicit.

participant is cued to the perspective of an engineering student charged with discovering the structure of a series of tubes contained inside a simple box. Marbles are dropped into two holes at the top of the box, and the student records which of two holes at the bottom of the box the marbles exit from. The student is given a diagram which shows one possibility for the internal tube structure, and is told to find out if the diagram is true or false. The participants' job is to indicate which holes must be checked to see if the diagram is wrong.

The concept of *gravity* was used to let the participant understand that the rule was both deterministic and unidirectional. A marble will enter through the top holes, fall through the series of tubes, then exit out of one of the bottom holes. This will always happen (deterministic), and will only happen in this way (unidirectional). The participants are also shown the relevance of looking for violations of the given rule – *"If* [the rule] *is wrong, then the diagram must be wrong"*.

The accompanying diagrams show (figure1) the box and the position of the holes, and (figure2) the diagram given to the engineering student to test. It should be noted that the "no picture" version of this task contained no diagrams, and all references to figures 1&2 were removed from the text (although, the participant *was told* that the character had been given a diagram, and that the highlighted rule would test the diagram).

The "*Automated Rail Controller*" scenario was another thematic-abstract task. The participant was cued to the position of a safety engineer charged with determining whether a manual override safety switch controlling an automated railway junction-controller is malfunctioning.

Participants' inherent understanding of simple electrical circuits was used to convey the concepts of "determinism" and "unidirectionality" in the tested rule. People know that closing a light switch will create an electric circuit which will "turn on the light". They know that this will happen every time (determinism), unless there is a malfunction. They also know that such a system is unidirectional; closing the switch will *cause* the light to turn on, but not vice-versa. Likewise, the participants should understand that the same deterministic and unidirectional criteria are relevant in the case of the manual over-ride circuit. The relevance of looking for violations of the tested rule is also expressed to the participants by telling them that malfunctions in the circuit could lead to serious rail accidents.

The accompanying diagram outlined the possible positions of the manual over-ride switch and it's relative effects on the rail-traffic-lights.

The "*Smithton Swamp*" scenario was the third of the thematic-abstract tasks constructed to the prescriptions of the theory of Task Understanding. The participant was cued to the perspective of an "expert in species-relocation" who is charged with finding out whether a certain rare species of frog would survive being relocated from it's natural habitat to another similar habitat. The critical concern is whether the frog can sustain itself on a diet other than marsh-flies, its only known food source.

The participants' inherent understanding of "food-chains" was used to convey the concept of "unidirectionality" in the tested rule. People understand that amphibians tend to eat insects, but that insects do not eat amphibians. The concept of "determinism" is supplied by suggesting that Smithton frogs only eat marsh-flies. The relevance of looking for violations of the tested rule is implied by suggesting that Smithton frogs would need to live off other insects if they were shifted to another swamp.

The accompanying diagram shows the two amphibian species, their known insect prey and the relationships showing which amphibian eats which species of insect.

Participants:

A chi-square was to be required for the analysis of "sex" and "picture" differences[1]. A power analysis revealed that a total of at least 117 participants would be required in order to gain enough statistical power (P=0.9, w=0.3, df=1, α=0.05). For counter balancing purposes, 144 participants were recruited; 72 males and 72 females.

As in Experiment 1a, participants were recruited by poster advertisement. All the participants went into a sweepstake for three prizes of NZ$75, NZ$50 and NZ$25. Recruitment took place at "*Student Job Search*" in the summer break. The same criterion were set for a person's inclusion in the experiment as for Experiment 1 (*see section 3.1*), with the addition that participants from Experiment 1 were excluded.

Apparatus:

Each participant was given a booklet containing four pages. The first page was an information and instruction sheet (shown in Appendix A2.3). The second page was a consent form for the participant to sign and date. The third and fourth pages were tasks from the two task conditions, one task per page. Participants were given one scenario from each condition. The orders of the task conditions (2 permutations) and the scenarios from each condition (9 combinations) were counterbalanced. Each combination was counterbalanced for males and females. In addition, half the participants received text scenarios with the accompanying pictures, while the other

---

[1] "Picture differences" are the differences in task performance between the tasks that had an accompanying picture or diagram, and the tasks that did not.

half received the text without the pictures, leaving a total of 72 combinations. This number was doubled to 144 by having two participants complete each combination.

The arrangement of the "cards" at the bottom of each copy of each scenario was randomised. The same arrangement of card orders was given to males and females. In other words, males and females received precisely the same scenarios, in the same orders, with the same arrangements of "cards" at the bottom of each scenario.

Procedure:

The procedure was identical to that in Experiment 1, except that data collect was shared by the author and a paid collaborator, who assisted when the author was unavailable.

A desk was set up at Student Job Search and a number of posters were displayed around the area advertising the study and the associated sweepstake. Interested people came to the desk where the experimenter briefly explained the task and ascertained whether the volunteer was qualified to do the experiment. If the volunteer passed the criterion, they wrote their name and contact details on a sheet of paper for the sweepstake and were given a task booklet.

The participant completed the task booklet in the Student Job Search. After all booklets were completed the sweepstake was drawn and the winners notified.

Results:

Nobody took longer than 15 minutes to complete the tasks. All participants circled at least one answer to every question. Two participants had to be replaced because they failed to return the booklets.

Table 4.1 shows the percentages of correct responses for each scenario, for sex of participant, and for each condition. The total numbers of correct responses for each category are shown in brackets.

Cochran's "Q" test for dependent samples was used to test for significant differences between the task conditions. The difference between the performance on the thematic-abstract condition (40%) and the SC cheating condition (50%) was found to be significant: $Q=5.49$, $df=1$, $p<0.05$. It should be noted that this difference, while significant, is fairly small. As in Experiment 1, the within-condition ranges are high compared to the between-condition difference. The difference between the two conditions is only 10%, while the range of the thematic-abstract scenarios is 31%, and the range of the SC cheating scenarios is 33%. So, while the difference between the

conditions is statistically significant, it is rather minimal when compared to the differences between the scenarios within each condition.

Table 4.1 Percentage of correct responses[1] (P&~Q) for Task Conditions, Individual Scenarios and Sex of Participant.

Task Conditions	Scenario	Scenario Performance (N=48)	Sex		Task Condition Performance (N=144)
			Males (N=72)	Females (N=72)	
Thematic-abstract	Engineering Diagram	38% (18)	49% (35)	31% (22)	40% (57)
	Rail Controller	25% (12)			
	Smithton Swamp	56% (27)			
SC cheating	Abstract SC	42% (20)	54% (39)	46% (33)	50% (72)
	Product Promotion	71% (34)			
	Security Guard	38% (18)			
	Means (M/F): N=144		51% (74)	38% (55)	

Analysing the individual scenarios from each condition, it is apparent that the major contributors to the between-condition difference are the high performance of the "Product Promotion" scenario (71%) and the poor performance of the "Automated Rail Controller" scenario (25%). This point is highlighted by the fact that if *either* scenario is removed from the analysis, the significant difference disappears:

Without "Product Promotion" scenario:
Thematic-abstract scenarios=35%, SC cheating=40%.
(N=96; Q=0.67; df=1; p= *not significant*).
Without "Automatic Rail Controller" scenario:
Thematic-abstract scenario=46%, SC cheating=48%.
(N=96; Q=0.04; df=1; p= *not significant*).

This loss of significance cannot be ascribed to a loss of statistical power as the number of subjects is still high (N=96). Rather, the between-conditions differences have all but disappeared (note the very low "Q" values).

[1] All results are rounded to the nearest full percentage.

Sex differences were also examined. Overall males (51%) performed better than females (38%). This was significant: (χ^2=5.04; N=288; df=1; p<0.05). However, the bulk of this variation is due to the performance disparity found between males (49%) and females (31%) on the Thematic-abstract task (χ^2=4.91, N=144, df=1, p<0.05). In fact, the poor performance by female participants on the Thematic-abstract condition seems to account for the bulk of the difference *between* the task conditions. The difference between females' performances across the two conditions is significant. But the difference between males' performances across the conditions is not significant.

Females:

Thematic-abstract (31%) vs. SC cheating (46%)

(Q=4.84, N=72, df=1, p<0.05).

Males:

Thematic-abstract (49%) vs. SC cheating (54%)

(Q=1.00, N=72, df=1, p= *not significant*).

Hence, there appears to be a significant "*Sex*" X "*Task condition*" interaction for females on the Thematic-abstract task.

Table 4.2 **Percentage of correct responses (P&~Q) for each scenario with respect to the presence or absence of an accompanying picture.**

Task Condition	Scenario	Picture (N=24)	No Picture (N=24)
Thematic-Abstract	Engineering Diagram	42% (10)	33% (8)
	Rail Controller	25% (6)	25% (6)
	Smithton Swamp	58% (14)	54% (13)
SC cheating	Abstract SC	42% (10)	42% (10)
	Product Promotion	71% (17)	71% (17)
	Security Guard	42% (10)	33% (8)
	Means: (N=144)	47% (67)	43% (62)

The effect of including the pictures/diagrams with the scenarios is summarised in Table 4.2. This shows the percentage of correct responses to each scenario for the instances where the pictures were included ("Picture") and not included ("No Picture"). The total number in each category is in brackets. The results show that the presence of

the pictures or diagrams had no significant effect on the percentage of correct responses (χ^2=0.35, N=288, df=1, p= *not significant*). No scenario differed by more than 2 correct responses between the "Picture"/"No Picture" categories.

Finally, an analysis of errors was completed for each scenario. Figure 4.1 gives the frequency of each response for the three Thematic-abstract scenarios.

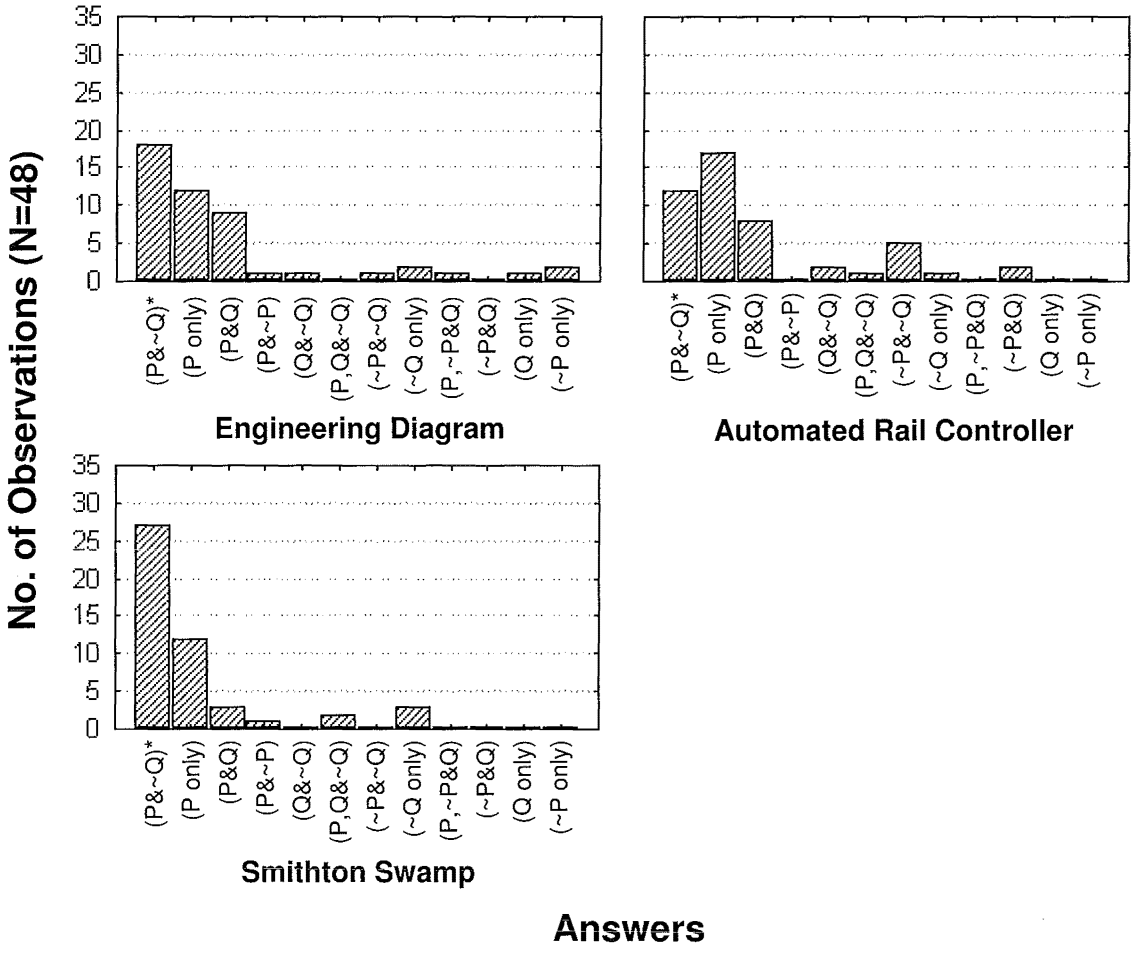


Figure 4.1: Graphs showing the pattern of responses to the three Thematic-abstract scenarios in Experiment 2. The correct answer is shown by *.

In all three thematic-abstract scenarios, the most frequent error is "**P only**", followed by "**P&Q**". As noted in Experiment 1, these are the most frequent errors found on the Wason Selection task, and are generally attributed to matching bias. No other error category is frequent enough to indicate any other particular biases on any of the scenarios.

Figure 4.2 gives the frequency of each response for the three SC cheating scenarios.

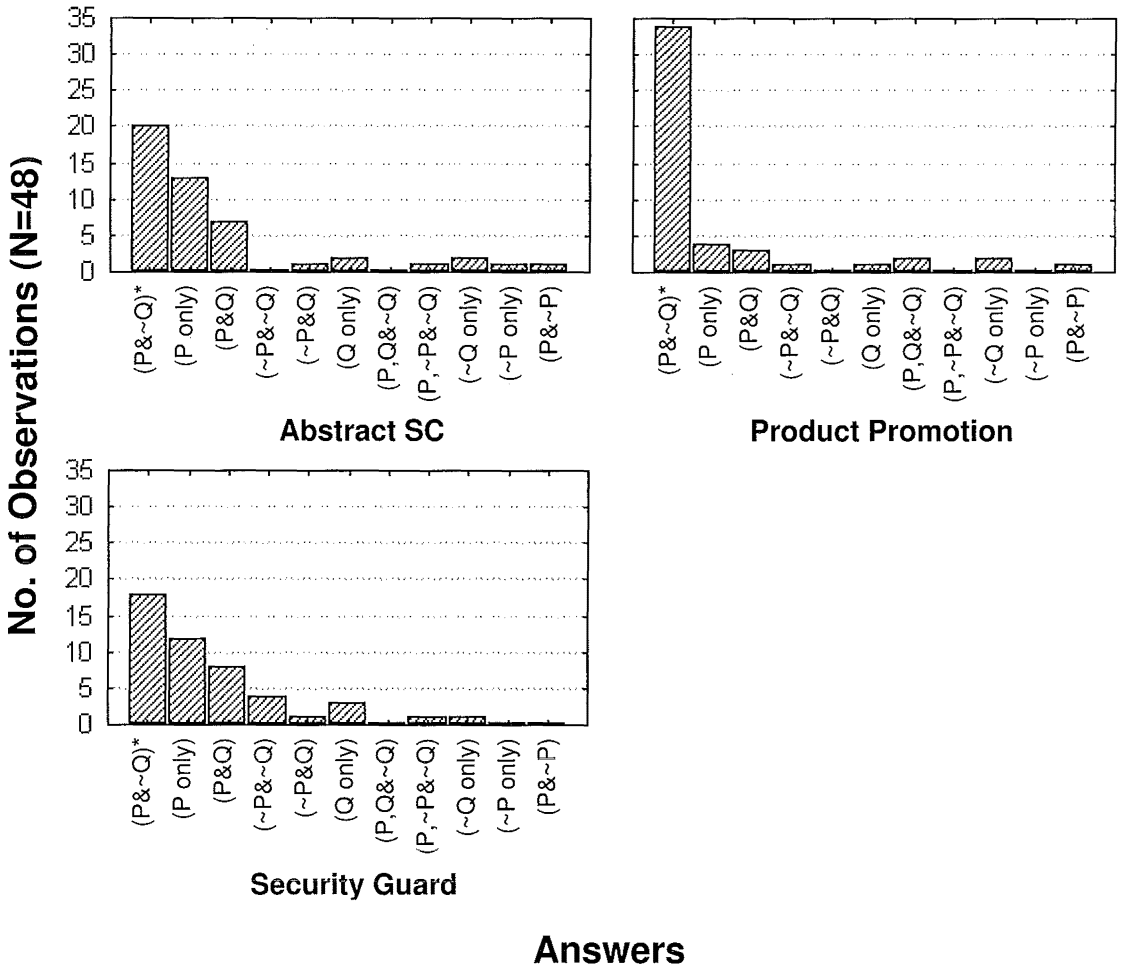


Figure 4.2: Graphs showing the pattern of responses to the three Social Contract cheating scenarios in Experiment 2. The correct answer is shown by *.

Once again, the most frequent error is "**P only**", followed by the "**P&Q**" category. None of the other error categories are frequent enough to indicate any unusual biases in any of the SC cheating scenarios.

Discussion:

The first main effect found in Experiment 2 was that the overall SC cheating condition performance was significantly higher than the overall performance of the Thematic–abstract condition. The putative conclusion to be drawn from Experiment 2 is that the prediction of Social Exchange theory – that SC cheating tasks would elicit higher rates of correct responding than the Thematic-abstract tasks – is upheld. However, a deeper analysis reveals that this conclusion maybe somewhat premature.

Firstly, this difference is small (10%). One of the most compelling arguments for Social Exchange theory has always been that the differences between social contract

tasks and abstract tasks have been very large (typically 50% or more). Therefore, it seemed reasonable to believe that a genetic predisposition to reason successfully in evolutionarily important situations (such as social exchanges) might be at the heart of this phenomenon. But when results, such as these, indicate that the performance benefits are low, the existence of a specialised evolved mechanism for dealing exclusively with social contracts becomes less credible.

Secondly, the main contributors to the between-condition difference were (a) the high success rate for the "Product Promotion" in the SC cheating condition, and (b) the low success rate for the "Automated Rail Controller" in the Thematic-abstract condition. The other four scenarios from both conditions produced very similar levels of success. The low level of responding of the "Automated Rail Controller" is particularly relevant. A post-hoc review of this scenario revealed that a mistake was made in the construction of the text. In a number of places, the instruction to look for *violations* of the rule was *implicit* ("*Your job... is to ensure that the manual over-ride switch is* ***working properly***" and "*Indicate only those card(s) that you definitely need to turn over to see if the manual over-ride switch is* ***working properly***"). In another place the participant was asked to *verify* the rule ("*...you are to check that the following rule is* ***true:***"), as opposed to *falsifying* the rule. Griggs and Jackson (1990) and Margolis (1987; in Evans et al., 1993) have found that these factors have an adverse impact on task performance. It also contravenes the third fundamental aspect of task understanding from Liberman & Klar's theory by reducing the perceived relevance of looking for a violation strategy. This mistake was not made in the construction of the other five scenarios, and may at least partially explain the relatively low level of performance for this task.

Finally, the overall difference in task performance between the conditions becomes less clear-cut when sex differences are considered. Only the females showed a significant performance difference between the task conditions. There was no significant performance difference between the conditions for the male participants. This is contrary to the prediction made by Social Exchange theory that the difference in performance of the task conditions should be the same for both sexes.

The poor performance of the females on the Thematic-abstract scenarios might be due to a kind of population-gender bias. Although all the scenarios avoided sexist language and gave no suggestion to the sex of the character who's perspective the participants were cued to, the nature of the tasks themselves might be gender biased. Both the "*Engineering Diagram*" and the "*Automated Rail Controller*" scenarios cued the participants to the positions of engineers. As noted in the general discussion to

Experiment 1, engineering and other similar technically oriented disciplines (such as physics, mathematics, and computer science) are courses primarily populated by males[1]. By this fact alone, the male sample from the tested population may contain a higher percentage of students from these technical domains. Therefore, some of the males from the sample may have a certain amount of domain expertise that would enhance task understanding[2]. This hypothesis is untested, however. What is needed is for the performance of people from these technical backgrounds to be tested against people from other backgrounds on these particular tasks. This information is not presently available, but could be a promising line for future research.

These three points suggest that the putative conclusion – that the Social Contract cheating condition out performs the Thematic-abstract condition – may be false. If this is the case, then this contradicts Social Exchange theory. If it is shown in further studies that no difference exists between performance on social contract and thematic-abstract tasks, then this will support the theory of Task Understanding.

The second main effect found in Experiment 2 was that the presence of a picture or diagram outlining the problem had no significant effect on task performance. This was contrary to the hypothesis that these pictures and diagrams should help task understanding and lead to better task performance. The conclusion is that this evidence fails to support the theory of Task Understanding.

However, in retrospect, it is probable that the pictures failed to fulfil the role that they were designed for. In terms of the three fundamental aspects of task understanding, the pictures only appear to convey the first one. The pictures suggest that the rules are *deterministic* and *unidirectional* (the first aspect), but fail to convey that **P&~Q** is the successful falsifying instance (this was left implicit in the diagrams) and that the participant should look for violations of the rule. This problem might be alleviated in the diagrams that use arrows to represent the structural relationships between the elements of the problem by adding a special "negative arrow" between the "**P**" and "**~Q**" elements which indicates that this relationship is *not allowed* by the rule. Again, this possibility has not yet been tested, but represents another possible line of further research.

---

[1] This is true at the University of Canterbury, in any case.

[2] Any improvement in performance is unlikely to come from *availability,* as it is unlikely that participants would have any experience with these specific problems.

Finally, special interest was taken in the performance of the "*Abstract Social Contract*" scenario (42%) which used *implicit* negatives in the cards. Jackson & Griggs (1990) found that "*Abstract Permission rules*" that used implicit negatives averaged only an 8% success rate. The results found here appear closer to "*Abstract Permission rules*" that used *explicit* negatives, which Jackson and Griggs found to be about 53% correct. Furthermore, the performance of the "Abstract Social Contract" scenario is about the same as that of the "Security Guard" SC cheating scenario (38%). This suggests that facilitation can occur on social contracts tasks without the need for a thematic context. This gives support to Social Exchange theory. However, because the scenarios also conform to the three fundamental aspects of task understanding, these results also support Liberman & Klar's theory. Although this result does not provide a discriminating test of the two theories, it is still an interesting result in it's own right. It shows a fundamental performance difference between the "*Abstract Social Contract*" and Cheng & Holyoaks' (1985) "*Abstract Permission rule*". However, if the text of the surrounding scenario to the abstract permission rule could be constructed to the prescriptions of Liberman & Klar's (1996) theory, then this *could* provide a discriminating test of Social Exchange theory and the theory of Task Understanding.

5 General Discussion:

5.1 Summary/Conclusions:

Experiments 1 and 2 addressed a variety of issues concerning the Wason Selection Task. In particular, three specific queries were investigated.

The first query addressed was "Is the cost/benefit relationship of the social contract necessary for facilitation?" Previous experiments had shown that the *detection of cheating* appeared to be necessary, but this could not be tested independently of the cost/benefit relationship of a social contract rule. Experiment 1 attempted to tease apart these factors by comparing performance on social contract cheating-detection tasks and non-social-contract *deceit-detection* tasks, which did not incorporate a cost/benefit relationship in the tested rule. The results were not straightforward. From one perspective, the results indicated that the cost/benefit relationship *was* necessary. Participants performed the Social Contract cheating-detection tasks significantly better than the Non-Social Contract deceit-detection tasks. However, the Non-Social contract deceit-detection tasks induced significantly better performance than a control Thematic-abstract task. Hence, some facilitation was apparent, even without the presence of the cost/benefit relationship.

These results were problematic for both Social Exchange theory and the general deceit-detection hypothesis. Social Exchange theory predicted that performance on the NSC deceit and the Thematic-abstract tasks should be equivalent, while the general deceit-detection hypothesis suggested that performance on the NSC deceit condition should be equivalent to that of the SC cheating condition. However, these results could be explained in terms of *task understanding*. Firstly, the NSC deceit tasks tended to be longer than the SC cheating tasks. Therefore, there might have been too much information in the NSC deceit tasks for the participants to successfully process. Also, the SC cheating tasks specifically asked participants to look for *violations* of the rule, while the NSC deceit tasks did not make this explicit. The perceived relevance of looking for violations of the tested rule is one of the fundamental aspects of Liberman & Klar's theory of Task Understanding.

The second query addressed was "Is it possible to construct abstract tasks that induce levels of performance comparable to that elicited by social contract tasks?" Experiment 2 tested the theory of Task Understanding against Social Exchange theory. Sets of Thematic-abstract scenarios were constructed to the prescriptions that the

theory of Task Understanding suggested should lead to good task performance. These were compared with SC cheating scenarios. The initial results supported Social Exchange theory. The Thematic-abstract scenarios *did* facilitate performance, but not to the same levels as SC cheating tasks. However, a more detailed analysis revealed that the differences between the conditions relied on the relatively poor performance of the females on the Thematic-abstract tasks. This might be accounted for by a possible population based gender-bias in the two lowest performing Thematic-abstract tasks. If this is true, then eliminating this gender-bias may eliminate the difference between the Thematic-abstract tasks and the SC cheating tasks.

In addition with this, the *between-condition diff*erence (10%) was substantially less than the *ranges between scenarios within each condition* (about 30%). This large variation in within-condition success rates is difficult to explain using Social Exchange theory, as its sole predictor of task performance is the presence or absence of cheater-detection in the scenario. Even if a cheater-detection algorithm *was* responsible for the small between condition difference, the large within-condition variance must be attributable to *some other factor or factors.* On the other hand, the theory of Task Understanding predicts that task performance will fluctuate with respect to how well the three fundamental aspects are understood. Therefore, although the initial results appeared to support Social Exchange theory, more detailed analysis appears to support the Theory of Task Understanding.

The third query addressed was "Does the inclusion of pictures or diagrams, which represent the task problem, increase task performance?" Half the tasks in Experiment 2 included accompanying pictures or diagrams which represented various elements of the scenario. It was hypothesised that these pictorial representations would reinforce the text, reduce ambiguity and provide additional methods of supplying the information in the problems. This should have lead to better task understanding, and therefore better performance. The results, however, showed that the presence or absence of such pictures had no significant effect on task performance. While this could be interpreted as a blow to the Theory of Task Understanding, it should be noted that the diagrams failed to adequately represent two of the three fundamental aspects of task understanding. So, even though these results did not find that pictorial representations improved task performance, future studies may yet prove fruitful in this area.

The results from Experiments 1 and 2 also unearthed a number of expected or auxiliary results. Firstly, both Experiments produced significant, if rather erratic, sex differences on task performance. Males tended to out perform females. These results were unexpected and have not been reported in the literature before (possibly because

nobody else has *looked* for them). While these differences could be due to population gender-bias in the cases of the Thematic-abstract tasks in Experiment 2, this seems unlikely to be the cause in Experiment 1. Whether this effect is real or simply a statistical aberration needs to be established with further testing.

Secondly, the results of Experiment 1 seemed to suggest that performance on some task-conditions might be dependent previously presented tasks. For instance, it appeared that the NSC deceit tasks in Experiment 1a elicited lower performance when they were preceded by a (poor performing) Thematic-abstract task. When the thematic-abstract tasks were removed from Experiment 1b, the overall performance of the NSC deceit tasks improved. This seems consistent with the predictions of the Theory of task Understanding: interpretation of one task should effect the interpretation of following tasks. However, this hypothesis was not tested directly, and further research could investigate this phenomenon in greater detail.

Finally, Experiment 2 investigated an auxiliary prediction that performance on the "*Abstract Social Contract*" scenario would be higher than that found by Jackson & Griggs (1990) for a similar "*abstract permission rule*". Both tasks used *implicit* negatives on the 'cards'. Previously, Jackson and Griggs found that performance was good (about 53%) on the "abstract permission rule" when *explicit* negatives[1] were used, but performance was poor (about 8%) when *implicit* negatives were used. The "*Abstract Social Contract*" scenario recorded 42% correct responses, which is more inline with the performance of Jackson & Grigg's participants on the explicit version of the "abstract permission rule" than the implicit version. This, too, is consistent with the theory of Task Understanding, because Jackson & Griggs did not construct their tasks to the prescriptions defined by the theory. However, this comparison has not been directly tested because Jackson & Griggs' experiment differs methodologically from the one described here (different populations and different presentation methods). Another experiment needs to be run which tests "Abstract Social Contract" and "Abstract permission rule" scenarios, with "explicit" and "implicit" negative cards counterbalanced across each condition.

---

[1] The use of explicit negatives has been found to reduce matching bias and increase logical performance (Evans, et al., 1993).

5.2 Theoretical Implications:

Although the series of experiments illustrated in this thesis do not provide clear-cut support for any theory of propositional reasoning, the results appear to favour Liberman & Klar's (1996) Theory of Task Understanding. This theory has a couple of advantages over its competitors.

The first advantage is that it is the most complete theory. It is a global theory that attempts to explain performance across all thematic domains by application of the same principles. The other theories (the domain-specific theories of "Social Exchange", Pragmatic Reasoning Schema" and the "Deceit-detection hypothesis", and the experientially based "availability Heuristic") are rather too specific to adequately explain task performance outside their domains.

This, in turn, provides a second advantage – one of *parsimony*. It is simpler to explain a given set of phenomena using one global theory, than it is to construct a number of separate theories to explain each phenomenon in the set. While the global theory might be more complicated than any of the separate theories, the combined effect of the separate theories may be far more complicated than the global theory. Such is the case with pragmatic reasoning schema. Although each pragmatic reasoning schema is, in itself, fairly basic, the shear number of schemas needed to explain every domain of facilitation may become very large[1].

Another area where the Theory of Task Understanding my be more parsimonious than it's competitors is in the complexity of cognitive mechanisms. The theory does not posit any new cognitive mechanisms (such as Social Exchange theory's "cheater-detection algorithm") or knowledge structures (such as pragmatic reasoning schemas). This makes the Theory of Task Understanding simpler because it does not rely on unproven evolutionary or developmental components.

---

[1] Although only two pragmatic reasoning schema have been popularised in the literature ("Permission schema" and "Obligation schema"), there are probably many other domains where facilitation occurs that do not conform to either of these schemas.

The main disadvantage of the Theory of Task Understanding is that it is essentially a theory designed *specifically* to explain performance on the *Wason Selection Task*. The selection task is simply one arbitrary tool for measuring simple propositional reasoning performance. Yet Liberman & Klar's theory has not currently been adapted for use in any other form. This is not to say that it could not be adapted, indeed it could (and should), but it currently has a very limited usage – particularly in 'the real world'. The basic tenants of the theory still seem valid when applied outside the selection task. It is possible that people *do* reason successfully when they *understand* a problem well enough. Also, the three fundamental aspects of task understanding, with a certain amount of reworking, may be used generically to predict peoples' performance on a range of propositional problems. With time, they may even provide ways of *improving* peoples' reasoning skills, or provide ways to ensure that mistakes relating to reasoning are *minimised* or *avoided*.

One way to increase the theory's generality is to incorporate it with other existing theories. In section 1.5 in was noted that the most successful theory across all realms of deductive reasoning was the *Mental Models* theory (Johnson-Laird & Byrne, 1991). The major problem with this theory was that it incorporates a final processing stage known as "*fleshing out the implicit model*". In this stage, a putative conclusion (constructed from the initial model) is tested against alternative possible models for instances of violation. Just how this "fleshing out" process is achieved is never specified by Johnson-Laird & Byrne. The Theory of Task Understanding may provide some light on this process. At the very least, it should describe *when* or *under what situations* the "fleshing out" process will lead to successful generation of a correct conclusion. The Mental Models theory, in turn, could provide the Theory of Task Understanding with a vehicle for application beyond the Wason Selection Task.

5.3 Future Research:

Many suggestions for areas of further research have already been articulated in section 5.1. However, broader issues relating to future research still remain.

Firstly, further research needs to be done on Liberman & Klar's theory of Task Understanding. In particular, the relative *importance* of the "*three fundamental aspects*" that Liberman and Klar present should be untangled. They place no particular rank or importance to each of the "aspects". Yet it appears to this researcher that the "third" aspect – "*the perceived relevance of looking for violations of the tested rule*" – may be

the most facilitatory. Remember, the most common single error on most versions of the Wason Selection task is not the failure to select the "**P**" card, or the erroneous selection of the "**~P**" or "**Q**" cards, but *the failure to select the "~**Q**" card.* This amounts to a failure to recognise that the rule must be *falsified.* The suggestion here, then, is that research should be conducted which independently manipulates the three fundamental aspects of task understanding. The differences in task performance should provide an estimate of the relative importance of each of the aspects. However, in practice this may be difficult to achieve. Manipulating one aspect will, in most cases, have some effect on the other two. For instance, manipulating the *'determinism' or 'direction'* ("aspect 1") of a tested rule may disrupt the perception of "**P&~Q**" as the relevant falsifying instance ("aspect 2").

Another possibility to test is the existence of *other* fundamental aspects of task understanding. Although Liberman & Klar suggest three aspects, this does not preclude the possibility of other aspects. In section 2.4, an argument was presented against Cosmides' "switched social contracts" by suggesting that many of these scenarios were *temporally inconsistent* with the tested rule. Hence, this might satisfy the conditions for inclusion as a fourth fundamental aspect – *the "temporally consistency" of the scenario.* "When the rule is temporally consistent with the text of the scenario, task understanding will be better, and performance will be higher". Again, this is an untested hypothesis – only further research will ascertain whether this factor is indeed relevant.

The final two suggestions for further research are not specifically concerned with the Theory of Task Understanding, or any particular theory. They are concerned with the whole field of research on deductive reasoning itself.

Firstly, there has been a strong tendency in the literature for researchers to promote their theories at the expense of competing theories. Often, however, the theories are not incompatible. A good example of this was Cosmides' attack on the availability heuristic. She claimed that because Social Exchange theory seemed to explain the results of her experiments better than availability, the availability heuristic should be rejected in favour of Social Exchange. Yet there is ample evidence in the literature (albeit somewhat anecdotal) that availability *does* seem to account for *some* of the facilitation and variation on the selection task. Indeed, it would be a somewhat surprising finding if it did not!

What is suggesting here, is that more integration of current theories needs to be achieved. It is unlikely that any current theory of reasoning will ever be able to explain *all* of the performance characteristics of the Wason Selection task or any other

deductive reasoning problem. However, an integration of many of these theories under a common framework may provide fruitful. In a sense, the Theory of Task Understanding is already a good example of this kind of integration. There are actually no new phenomena described in Liberman & Klar's work. What they have done is taken a number of factors which were known to have an influence on task performance and provided a framework that integrates them into a coherent theory.

Related to this is the ever-present dichotomy between global theories of reasoning and domain-specific theories. Proponents of theories from both realms appear to argue that reasoning must be governed by either one field or the other[1]. This is almost certainly wrong. Certain global theories such as "Task Understanding" and "Mental Models" indicate that *most* instances of reasoning can be explained by appealing to the same mechanisms or phenomena. However, this is only the general case; domain-specific reasoning may still explain certain instances *better* than global theories. In particular, the effect of domain-specific knowledge seems to induce better performance within the set domain. And given what is known about the nature of *expertise* and *reasoning by analogy* (Anderson, 1990) it is difficult to believe that *schema*-based theories do not have some part to play in the future of deductive reasoning theory.

The last point to be made concerns the extraordinary amount of research concentrated on the Wason Selection Task. This task is simply *one* tool which can be use to gauge propositional deductive reasoning performance. Yet it accounts for the vast majority of research in this field. The reasons for this are obvious: it provides a concurrent test of all four propositional outcomes ("**P**", "**~P**", "**Q**" and "**~Q**") and lets the experimenter manipulate various aspects of the problem via the instructions and scenario. It is a very useful tool.

The problem, however, is that it detracts from other areas of propositional reasoning. The danger is that while we may learn a considerable amount about the Wason Selection Task, we may learn very little about propositional reasoning itself! Already a number of theories have been presented which, in their current form, only attempt to explain the selection task (for example, the "Theory of Task Understanding", "Pragmatic Reasoning Schema", and "Social Exchange theory"). The challenge, then, is to adapt these theories to other areas, and to design new methods and tools to test them.

---

[1] For instance, Cosmides (1989) and Cosmides & Tooby (1987) suggest that global theories cannot work because global mechanisms could never evolve. Rather, a large number of purpose-specific mechanisms have evolved which we naively identity as a single mechanism or phenomena.

The final goal should be to take these theories out of the realm of theoretical cognition, and apply them to real world settings. Once deductive problem areas can be identified, and techniques advanced to solve the problems, technology can be developed to reduce problems and mistakes associated with deductive reasoning. A wealth of research in the last two decades has resulted in an explosion of human-factors technology. Cognitive technology comprises a relatively new branch of this field. With time, it is hoped, the theories described in this thesis may become the bases of technology and culture which will be integrated into our everyday lives.

References:

Anderson, J. R. (1990). *Cognitive psychology and its implications (3rd ed.).* New York: W. H. Freeman

Baddeley, A. D. (1986). Working Memory. Oxford: Claredon Press.

Braine, M. D. S., & O'Brien, D. P. (1991). A theory of If: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review, 98,* 182-203.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391-416.

Cheng, P. W., & Holyoak, K. J. (1989). On the natural selection of reasoning theories. *Cognition, 33,* 285-313.

Cheng, P. W., Holyoak, K. J. Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic verses syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293-328.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, N.J.: Lawrence Erlbaum.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-276.

Cosmides, L. & Tooby, J. (1987). From evolution to behaviour: Evolutionary psychology as the missing link. In J. Dupré (Ed.). *The latest on the best: Essays on evolution and optimality.* Cambridge, MA: MIT Press.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning.* London: Routledge & Kegan Paul.

Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75,* 451-468.

Evans, J. St. B. T. (1995). Relevance and reasoning. In S. E. Newstead & J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason*. Hove, UK: Lawrence Erlbaum.

Evans, J. St. B. T., & Lynch, J. S. (1973). Matching bias in the selection task. *British Journal of Psychology, 64,* 391-397.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction*. Hove, U.K.: Lawrence Erlbaum.

Gigerenzer, G., & Hug, K. (1992). Social contracts, cheating, and perspective change. *Cognition, 43,* 127-171.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic materials effect in the Wason selection task. *British Journal of Psychology, 73,* 407-420.

Hardman, D. K. (1996). *Adaptive decision making and the Wason selection task*. Unpublished manuscript, City University, London.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology, 7,* 1-52.

Jackson, S. L., & Griggs, R. A. (1988). Education and the selection task. *Bulletin of the Psychonomic Society, 26,* 327-330.

Jackson, S. L., & Griggs, R. A. (1990). The elusive pragmatic reasoning schema effect. *Quarterly Journal of Experimental Psychology, 42a,* 353-374.

Johnson-Laird, P. N. (1993). *Human and machine thinking.* Hillsdale, N.J.: Lawrence Erlbaum.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hove, U.K.: Lawrence Erlbaum.

Kirby, N. K. (1994a). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition, 51,* 1-28.

Kirby, N. K. (1994b). False alarm: a reply to Over and Evans. *Cognition, 52,* 245-250.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94,* 211-228.

Liberman, N., & Klar, Y. (1996). Hypothesis testing in Wason's selection task: social exchange cheating detection or task understanding. *Cognition, 58,* 127-156.

Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition, 39,* 85-105.

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information.* San Francisco: Freeman.

Oakhill, J.V., & Johnson-Laird, P.N. (1985). The effects of belief on the spontaneous production of syllogistic conclusions. *Quarterly Journal of Experimental Psychology, 37A,* 553-369.

Over, D. E., & Evans, J. St. B. T. (1994). Hits and misses: Kirby on the selection task. *Cognition, 52,* 235-243.

Over, D. E., & Manktelow, K. I. (1993). Rationality, utility and deontic reasoning. In K. I. Manktelow & D. E. Over (Eds.). *Rationality: Psychological and philosophical perspectives.* London: Routledge.

Politzer, G., & Nguyen-Xuan, A. (1992). Reasoning about conditional promises and warnings: Darwinian algorithms, mental models, relevance judgement or pragmatic reasoning schema? *Quarterly Journal of Experimental Psychology, 44A (3),* 401-421.

Pollard, P. (1982). Human reasoning: Some possible effects of availability. *Cognition, 12,* 65-96.

Rips, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review, 90,* 38-71.

Rips, L. J. (1989). The psychology of knights and knaves. *Cognition, 31*, 85-116.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge, MA: MIT Press.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: M^{c}Graw-Hill.

Taplin, J. E., & Staudenmayer, H. (1973). Interpretation of abstract conditional sentences in deductive reasoning. *Journal of Verbal Learning and Verbal Behaviour, 12,* 530-542.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46,* 35-57.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 4,* 207-232.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology, Vol. 1.* Harmondsworth, UK: Penguin.

Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content.* London: Batsford.

Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioural sciences. *Behavioural and Brain Sciences, 17,* 585-654.

Appendix I – Previous tasks

A1.1 Cosmides' Unfamiliar Social Contracts:

(1) Cassava Root Law (Social Law):

You are a Kaluame, a member of a Polynesian culture found only on Maku Island in the Pacific. The Kaluame have many strict laws which must be enforced, and the elders have entrusted you with enforcing them. To fail would disgrace you and your family.

Among the Kaluame, when a man marries, he gets a tattoo on his face; only married men have tattoos on their faces. A facial tattoo means that a man is married, an unmarked face means that a man is a bachelor.

Cassava root is a powerful aphrodisiac – it makes the man who eats it irresistible to women. Moreover, it is delicious and nutritious – and very scarce.

Unlike cassava root, molo nuts are very common, but they are poor eating – molo nuts taste bad, they are not very nutritious, and they have no other interesting "medicinal" properties.

Although everyone craves cassava root, eating it is a privilege that your people closely ration. You are a very sensual people, even without the aphrodisiacal properties of cassava root, but you have strict sexual mores. The elders strongly disapprove of sexual relations between unmarried people, and particularly distrust the motive and intentions of bachelors.

Therefore, the elders have made laws governing rationing privileges. The one you have been entrusted to enforce is as follows:

"If a man eats cassava root, then he must have a tattoo on his face".

Cassava root is so powerful an aphrodisiac, that many men are tempted to cheat on this law whenever the elders are not looking. The cards below have information about four young Kaluame men sitting in a temporary camp; there are no elders around. A tray filled with cassava root and molo nuts has just been left for them. Each card represents one man. One side of a card tells whether or not the man has a tattoo on his face.

Your job is to catch men whose sexual desires might tempt them to break the law – if any get past you, you and your family will be disgraced. Indicate only those card(s) you definitely need to turn over **to see if any of these Kaluame men are breaking the law.**

The cards read: "eats cassava root", "eats molo nuts", "tattoo", and "no tattoo".

(2) Duiker meat Law (Social Law):

You are an anthropologist studying the Namka, a hunter-gather culture living in the deserts of south-west Africa. You are particularly interested in whether Namka boys obey the laws of their people.

Every full moon there is a special feast in which a duiker – a small antelope – is slaughtered and eaten. Duiker meat is quite scarce and delicious – a real treat. Eating duiker meat is a privilege that must be earned.

For boys, this privilege is governed by the following law:

"If you eat duiker meat, then you have found an ostrich shell".

Finding ostrich eggshells is a sophisticated and difficult task which takes a boy years to learn. Having found an ostrich eggshell on your own is therefore a sign that you have mastered the most difficult skills of hunting. For the Namka, it represents a boy's transition to manhood.

You wonder if Namka boys cheat on this law when nobody is looking. You decide to hide behind some bushes and watch. During the course of the feast of the full moon, you see four different boys approach the roasted duiker while no one else is looking.

The cards below have information about these four boys. Each card represents one boy. One side of a card tells whether a boy has ever found an ostrich eggshell, and the other side of the card tells whether that boy took any roasted duiker meat.

The smell of the roasting duiker is truly tempting to the boys. You want to know if any of them cheated on the law. Indicate only those card(s) you definitely need to turn over **to see if any of these boys have broken the law.**

The cards read: "eats some duiker meat", "does not eat any duiker meat", "has found an ostrich eggshell", and "has never found an ostrich eggshell".

(3) Cassava Root Contract (Private Exchange):

You are and anthropologist studying the Kaluame, a Polynesian people who live in small, warring bands on Maku Island in the Pacific. You are interested in how Kaluame "big men" – chieftains – wield power.

"Big Kiku" is a Kaluame big man who is known for his ruthlessness. As a sign of loyalty, he makes his own "subjects" put a tattoo on their face. Members of other Kaluame bands never have facial tattoos. Big Kiku has made so many enemies in other Kaluame bands, that being caught in another village with a facial tattoo is, quite literally, the kiss of death.

Four men from different bands stumble into Big Kiku's village, starving and desperate. They have been kicked out of their respective villages for various misdeeds, and have come to Big Kiku because they need food badly. Big Kiku offers each of them the following deal:

"If you get a tattoo on your face, then I'll give you cassava root."

Cassava root is a very sustaining food which Big Kiku's people cultivate. The four men are very hungry, so they agree to Big Kiku's deal. Big Kiku says that the tattoos must be in place tonight, but that the cassava root will not be available until the following morning.

You learn that Big Kiku hates some of these men for betraying him to his enemies. You suspect he will cheat and betray some of them. Thus, this is a perfect opportunity for you to see first hand how Big Kiku wields his power. The cards below have information about the fates of the four men. Each card represents one man. On side of each card tells whether the man went through with the facial tattoo that evening and the other side tells whether or not Big Kiku gave that man cassava root the next day.

Did Big Kiku get away with cheating any of these four men? Indicate only those card(s) you definitely need to turn over **to see if Big Kiku has broken his word to any of these four men.**

The cards read: "got the tattoo", "no tattoo", "Big Kiku gave him cassava root", and "Big Kiku gave him nothing".

(4) Duiker meat Contract (Private Exchange)

The Namka are a hunter-gather people who live in small bands in the deserts of south-west Africa. You are an anthropologist interested in whether members of different Namka bands can trust each other.

Bo is a crafty old Namka man in the band you are studying. He is always accidentally breaking his ostrich eggshell and would like to "stockpile" some – the Namka use ostrich eggshells as canteens because they are light and hold lots of water. He sees his opportunity when four men from a neighbouring band stumble into camp one morning.

The four men have been on a long and unsuccessful hunting expedition. They are hungry, and they want to be able to bring back meat to their families. Bo approaches each man privately and offers him the following deal:

"If you give me your ostrich eggshell, then I'll give you duiker meat."

Bo explains that his wife is skinning the duikers today, and that they won't be ready until tomorrow. However, he will need the eggshell by this evening for his son, who is leaving tonight on a week long hunting expedition. Each man accepts Bo's offer, and agrees to meet him alone in a secluded spot tomorrow to consummate the deal.

You find this deal very interesting, because you happen to know that Bo, who is a rather unscrupulous character to being with, has very little duiker meat and a large family to feed. It is perfectly possible that he will cheat some of these men. You decide to "spy" on Bo and see.

The cards below have information about the four deals Bo made with these four men. What happened in one deal had no effect on the outcome of any other deal. Each card represents one man. On side of a card tells whether or not the man gave his ostrich eggshell to Bo that evening, and the other side of the card tells whether or not Bo gave that man duiker meat the next day.

Did Bo get away cheating any of these four men? Indicate only those card(s) you definitely need to turn over **to see if Bo has broken his word to any of these four men.**

The cards read: " he gave his ostrich eggshell to Bo", "He gave Bo nothing", "Bo gave him duiker meat", and "Bo gave him nothing".

A1.2 Cosmides' Unfamiliar Descriptive Tasks:

(1) Descriptive Cassava Root Problem:

You are an anthropologist studying the Kaluame people, a Polynesian culture found only on Maku Island in the Pacific. Before leaving for Maku Island you read a report that says that some Kaluame men have tattoos on their faces, and that they eat either cassava root or molo nuts, but not both. The author of the report, who did not speak the language, said the following relation seemed to hold:

"If a man eats cassava root, then he must have a tattoo on his face."

You decide to investigate your college's peculiar claim. When you arrive on Maku Island, you learn that the cassava root is a starchy staple food found on the south end of the island. Molo nuts are very high in protein, and grow on molo trees, which are primarily found on island's north shore.

You learn that bachelors live primarily on the north shore, but when men marry, they usually move to the south end of the island. When a Kaluame man marries, he gets a tattoo on his face; only married men have tattoos on their faces. A facial tattoo means that a man is married, an unmarked face means the man is a bachelor. Perhaps men are simply eating foods which are most available to them.

The cards below have information about four Kaluame men sitting in a temporary camp at the center of the island. Each man is eating either cassava root or molo nuts that he has brought from home. Each card represents one man. One side of a card tells which food a man is eating and the other side of the card tells whether or not the man has a tattoo on his face.

The rule laid out by your college may not be true; you want to see for yourself. Indicate only those card(s) you definitely need to turn over **to see if any of these Kaluame men are breaking the rule.**

The cards read: "eats cassava root", "eats molo nuts", "tattoo", and "no tattoo".

(2) *Descriptive Duiker meat Problem:*

You are an anthropologist studying the Namka, a hunter-gatherer culture in the deserts of south-west Africa. Over and over again, you hear various Namka repeat the following saying:

"If you eat duiker meat, then you have found an ostrich eggshell."

Duikers are small antelopes found in the eastern part of the Namka's home range. Both duiker meat and ostrich eggshells are sought by the Namka: they eat the meat and use the eggshells as canteens because they are light and hold lots of water. Furthermore, duikers frequently feed on ostrich eggs. As an anthropologist, you don't know if this saying is metaphorical, referring, for example, to clan territories or ritual practices, or if the saying reflects a real relationship the Namka use to guide their foraging behaviour. Does it mean that if you find the first you find the second? This is what you are trying to find out.

Is it fact or folklore? Do the Namka *mean* eggshells and duiker meat, or are these things merely symbols for something else entirely? Unfortunately, you don't know their language well enough to ask them. So you decide to investigate whether the rule stated in this saying has any *factual* basis.

Many species of bird populate the area, and in your wanderings you have come across several caches of eggs of various sorts. The cards below have information about four different locations with egg caches. Each card represents one location, and each location has the tracks of one mammal associated with it. One side of a card tells what kind of eggshell you found at a location, and the other side of the card tells which mammal's tracks you found there.

Perhaps the Namka's saying has no factual basis. Indicate only those card(s) you definitely need to turn over **to see if your finds at any of these locations violates the rule expressed in the Namka's saying.**

The cards read: "duiker", "weasel", "ostrich eggshell", and "quail eggshell".

A1.3 Cosmides' Familiar Descriptive Tasks:

(1) Boston/Arlington Transport Problem:

Part of your new job for the City of Cambridge is to study the demographics of transportation. You read a previously done report on the habits of Cambridge residents which reads:

"If a person goes to Boston, then he takes the subway."

The cards below have information about four Cambridge residents. Each cards represents one person. One side of a card tells where a person went and the other side of the card tells how that person got there.

Indicate only those card(s) you definitely need to turn over **to see if any of these people violate this rule.**

The cards read: "Boston", "Arlington", "cab" and "subway".

Note: Cosmides used four different versions of this task. The items in the rule were substituted between "Boston/subway", "Boston/cab", "Arlington/subway" and "Arlington/cab"

A1.4 Cosmides' Thematic-Abstract Tasks:

(1) Cosmides' Abstract Problem (Document coding task):

Part of your new clerical job at the local high school is to make sure that student documents have been processed correctly. Your job is to make sure the documents conform to the following alphanumeric rule:

"If a person has a 'D' rating, then his documents must be marked code 3."

You suspect the secretary you replaced did not categorise the students' documents correctly. The cards below have information about the documents of four people who are enrolled at this high school. Each card represents one person. One side of a card tells a person's letter rating and the other side of the card tells that person's number code.

Indicate only those card(s) you definitely need to turn over **if the documents of any of these people violates this rule.**

The cards read: "D", "F", "3" and "7".

Note: Cosmides used two different versions of this task. The items in the rule were substituted between "D & 3" and "C & 2".

A1.5 Cosmides' Culturally Familiar Social Contract Tasks:

(1) Grover High School problem.

You supervise four women who volunteered to help at the local Board of Education. When you came into work today, you found the place a-buzz with rumor and innuendo. Your volunteers were supposed to follow certain rules for assigning students from various towns to the appropriate school district. Each volunteer is the mother of a teenager who is about to enter high school, and each processed their own child's documents. So, now rumors are flying that your volunteers cheated on the rules when assigning their own children to a school. Here is the situation:

Students are to be assigned either to Grover High School, located in the Grover City, or to Hanover High School, located in the town of Hanover. Grover High is a great school with excellent record for getting students placed in good colleges. In contrast, Hanover High is a mediocre school with poor teachers and decrepit facilities.

The reason the schools are so different is due to how willing the parents of each community are to financially support their schools through taxes. Although both communities are equally prosperous, the parents of Grover have always cared about the quality of their schools, including Grover High, and have been willing to pay extra for it. In contrast, the parents in the neighboring

suburbs of Hanover and Belmont never wanted to spend the money, and have opposed any taxes to improve Hanover High.

The Board of Education took these factors into account when it created rules to determine which school a student is to be assigned to. The most important of these rules is:

"If a student is to be assigned to Grover High School, then that student must live in Grover City".

Your volunteers were supposed to follow this rule when processing *all* students documents - including the documents of their own children! You must find out if the rumors are true: did any of your volunteers cheat on this rule when it came to processing their own children's documents?

The cards below have information about the documents of the four volunteers' children. Each card represents the child of one volunteer. One side of a card tells which school the volunteer assigned their son or daughter to, and the other side of the card tells what town that student lives in.

Most parents want to get their children the best education possible, however, some are not willing to pay for it. It is easy to imagine that your volunteers, being ambitious parents, might have been tempted to cheat on the rule. Indicate only those card(s) you definitely need to turn over **to see if the documents of any of these students violate the rule**.

The cards read: "Grover High School", "Hanover High School", "Grover City" and "town of Hanover".

A1.6 Cosmides' Non-Social-Contract Permission Tasks:

(1) Grover High School problem (culturally familiar).

The secretary you replaced at the local Board of Education may have made some mistakes when she processed student documents. It is important that certain rules for assigning students from various towns to the appropriate school district is followed, because the population statistics they provide allow the Board of Education to decide how many teachers need to be assigned to each school. If the rules are not followed, some schools could end up with too many teachers, and other schools with too few.

Students are to be assigned to either Grover High School or to Hanover High School.

Some students live in the town of Grover City, some live in Hanover, and some live in Belmont. There are rules that determine which school a student is to be assigned to; the most important of these is:

"If a student is to be assigned to Grover High School, then that student must live in Grover City."

Shortly before she retired, the secretary you replaced was supposed to sort through the documents that specify what town the students' live in, and make school assignments according to this. She was a sweet old lady who had become rather absent-minded, and who often made mistakes when categorizing student documents.

The cards below have information about the documents of four students. Each card represents one student. One side of a card tells what school the retired secretary assigned the student to, and the other side of the card tells what town that student lives in.

You suspect the retired secretary may have inadvertently categorized some of the students' documents incorrectly, so you decide to see for yourself whether she ever violated the rule. Indicate only those card(s) you definitely need to turn over **to see if the documents of any of these students violate the rule.**

The cards read: "Grover High School", "Hanover High School", "Grover City" and "Town of Hanover".

(2) Cassava Root problem (culturally unfamiliar).

You are an anthropologist studying the Kaluame people, a Polynesian culture found only on Maku Island in the Pacific. The Kaluame people are divided into two great clams: the Napali, who distinguish themselves by getting tattoos on their faces when they are children, and the Kaloi, who have no facial tattoos. Members of the Napali and Kaloi clans live together in peace and friendship. Important matters are decided by a group of Kaluame elders, half of whom are Napali, and the other half, Kaloi.

Twenty years ago, when you first started studying the Kaluame, the elders became concerned about Maku Island's dwindling resources. The Kaluame do not cultivate food. Instead, they gather food that grows wild. They have two stable foods, which are equally tasty and nutritious: cassava root, a tubor which can grow only on the south end of Maku Island, and molo nuts, from molo trees which can only grow on the island's north shore. The problem is that both cassava root and molo nuts are in short supply, because the Kaluame population has been growing. If all the Kaluame lived on the south end of Maku Island, they would surely exhaust the supply of cassava root; if they all lived on Maku Island's north shore, they would surely exhaust the supply of molo nuts. The elders want their people to live in a balance with nature; they do not want to cause the extinction of either source or food, as this could eventually lead to the extinction of the Kaluame themselves.

Therefore, the elders decided to divide the Kaluame people in half, so that, roughly, one clan would live where the cassava root grows, and eat only the cassava root, and one clan would live where the cassava molo nuts grow, and eat only molo nuts. That way, neither food source would be overwhelmed by too many people, and both clans would be well nourished. Everyone agreed that this was a good plan. The only problem was that one clan had more people than the other, so 10% of the larger clan were asked to live and eat with the smaller clan. They gladly agreed.

The elders expressed the law governing eating arrangements thus:

"If a man eats cassava root, then he must have a tattoo on his face."

When you left Maku Island, everyone was happy observing this law, hoping that the cassava plants and molo nut trees would flourish as a result, so the next generation would not have to worry about such things.

That was 20 years ago. Now you are returning to Maku Island to continue your study of the Kaluame, and you wonder whether the plan worked. Because the Kaluame are such law-abiding people, the best way to see if the law is still in effect is to watch what the Kaluame are eating; if any of them are breaking it, then it is because the plants flourished and the elders repealed the law.

The cards below have information about four Kaluame men sitting in a temporary camp at the canter of the Island. Each man is eating either cassava root or molo nuts, which he brought with him from home. Each card represents one man. One side of each card tells which food the man is eating, and the other side of the card tells whether or not the man has a tattoo on his face.

The elders' law may no longer be in effect; the best way to tell is to see whether any of the Kaluame men are breaking it. Indicate only those card(s) you definitely need to turn over **to see if any of these Kaluame men are breaking the law.**

The cards read: "Eats cassava root", "eats molo nuts", "tattoo", "no tattoo".

(3) Duiker Meat problem (culturally unfamiliar).

You are an anthropologist studying the Namka, a hunter-gatherer culture living in the desert of southwest Africa. The Namka are divided into great clans: the Bakas and the Heronas. Although the two clans are quite similar, it is easy to tell the Bakas from the Heronas due to a minor cultural quirk. As children, members of the Baka clan become adept at the secret art of finding ostrich eggshells, which they use as canteens because they are strong and hold lots of water. Herona children, however, become adept at the secret art of making canteens from goat-skins; Herona children never learn how to find ostrich eggshells, just as Baka children never learn how to make goatskin canteens. Thus you can always tell the clans apart by seeing what kind of canteen a man has strapped to his side: a Baka carries an ostrich eggshell canteen, whereas a Herona carries a goatskin canteen. Members of the Baka and Herona clans live together in peace and friendship. Important matters are decided by a group of Namka elders, half of whom are Baka, and the other half, Herona.

Twenty years ago, when you fist started studying the Namka, the elders became concerned about the desert's dwindling animal population. The Namka hunt to get meat, and they particularly rely on two different species of small antelopes: duikers and gazelles. Duikers and

gazelles are equally tasty and nutritious, and equally easy to hunt. Duikers are found only in the eastern part of the Namka's home range, whereas gazelles are found only in the western part. The problem is that both duikers and gazelles are in short supply, because the Namka population has been growing. If all the Namka lived and hunted in the eastern half of the home range, they would surely exhaust the supply of duikers; if they all lived and hunted in the western half, they would surely exhaust the supply of gazelles. The elders want their people to live in a balance with nature; they do not want to cause the extinction of either food source, as this could eventually lead to the extinction of the Namka themselves.

Therefore, the elders decided to divide the Namka people in half, so that, roughly, one clan would live where the duikers roam, and eat only duiker meat, and one clan would live where the gazelles roam, and eat only gazelle meat. That way, neither food source would be overwhelmed by too many people, and both clans would be well nourished. Everyone agreed that this was a good plan. The only problem was that one clan had more people than the other, so 10% of the larger clan were asked to live, hunt, and eat with the smaller clan. They gladly agreed.

The elders expressed the law governing eating arrangements thus:

"If you eat duiker meat, then you have found an ostrich eggshell."

When you left the Namka, everyone was happily observing this law, hoping that the duikers and gazelles would flourish as a result, so that the next generation would not have to worry about such things.

That was 20 years ago. Now you are returning to southwest Africa to continue your study of the Namka, and you wonder whether the plan worked. Because the Namka are such law-abiding people, the best way to see if the law is still in effect is to watch what the Namka are eating; if any of them are breaking it, then it must be because the animals flourished and the elders repealed the law.

The cards below have information about four Namka men sitting in a temporary camp at the center of the home range; you can tell which clan each is from by their canteens. Each man is eating duiker meat or gazelle meat, which he brought with him from home. Each card represents one man. One side of each card tells which food a man is eating, and the other side of the card tells whether or not the man has ever found an ostrich eggshell.

The elders' law may no longer be in effect; the best way to tell is to see whether any of the Namka men are breaking it. Indicate only those card(s) you definitely need to turn over **to see if any of these Namka men are breaking the law.**

The cards read: "Eats duiker meat", "eats gazelle meat", "has found an ostrich eggshell" and "has never found an ostrich eggshell".

A1.7 Gigerenzer & Hugs' cheating/no-cheating social contract scenarios:

(1) The Overnight Problem[1]:

"The overnight rule ... is familiar to generations of mountain hikers in the Alps. Two context stories were used. The first explained that there is a cabin at high altitude in the Swiss Alps, which serves hikers as an overnight shelter. Since it is cold and firewood is not otherwise available at this altitude, the rule is that each hiker who stays overnight has to carry along his/her share of wood. There are rumours that the rule is not always followed. The subjects were cued into the perspective of a guard who checks whether any one of four hikers has violated the rule. The four hikers were represented by four cards that read "stays overnight in the cabin", "carried no wood", "carried wood", and "does not stay over night in the cabin". We refer to this version of the overnight problem as the *cheating version*.

In the *no-cheating version* the subjects were cued into the perspective of a member of the German Alpine Association who visits the Swiss cabin and tries to find out how the local Swiss Alpine Club runs this cabin. He [*sic*] observes people bringing wood to the cabin, and a friend suggests the familiar overnight rule as an explanation. The context story also mentions an alternative explanation: rather than the hikers, the members of the Swiss Alpine Club, who do not stay overnight, might carry the wood. The task of the subject was to check four persons (the same four cards) in order to find out whether anyone had violated the overnight rule suggested by the friend."

---

[1] Gigerenzer & Hug did not give the full text of their scenarios, rather, they gave overviews. The text reproduced here is the same as they gave in the paper.

A1.8 Liberman & Klar's unconfounded cheating/no-cheating social contract scenarios:

(1) *The Overnight Problem[2], unconfounded cheating version:*

"This version added to the original cheating version that according to the hiker's tradition, a hiker may deviate from the rule if he goes for an especially long track on the day after. In that case, he or she has to ask another hiker, who didn't sleep in the cabin, to bring wood. According to the tradition, this is suppose to be done for free, as a sign of good will and friendship among hikers. Th Alpine association suspects that some hikers take money for bringing wood. They consider this to be a rude violation of the hiker's tradition, and decide to put an end to the mess. This version suggests two deviations from normal interpretation. First, the rule is likely to be understood as implying that everyone who uses wood has to take care of the wood supply, but not necessarily bring it himself. Thus, the rule in its exact phrasing can be interpreted as non-deterministic. Secondly, the story suggests the possibility of someone who didn't stay overnight in the cabin and brought wood (not-P & Q) as a possible violation. This is suggested in addition to the original P & not-Q violating instance. Thus, the violating combination is no longer as clear as in the original cheating version, although the problem remains that of cheating detection an the original perspective is not changed."

(2) *The Overnight Problem, unconfounded no-cheating version:*

"This version stated that the rule is strictly obeyed by the hikers. Only guides who are members of the Alpine Association do not have to bring wood after staying overnight in the cabin. The guides have to announce their coming a few days in advance and the association takes care of the wood supply. Subjects were cued to the perspective of a hiker that looks for professional help of a guide for a friend who is stuck in the mountains. A list of visitors is available in the cabin, with information about were each visitor is to be found, whether he or she stayed overnight in the cabin and whether he or she brought wood. The nature of the violation that should be looked for is clear, but this violation does not involve cheating."

[2] Like Gigerenzer & Hug, Liberman & Klar only gave overviews of their tasks. The text presented here is reproduced from those in their paper.

Appendix II

Appendix A2.1

Tasks for Experiment 1

In all examples "card a"= **P**,
"card b"=**~P**
"card c"= **Q**
"card d"=**~Q**.

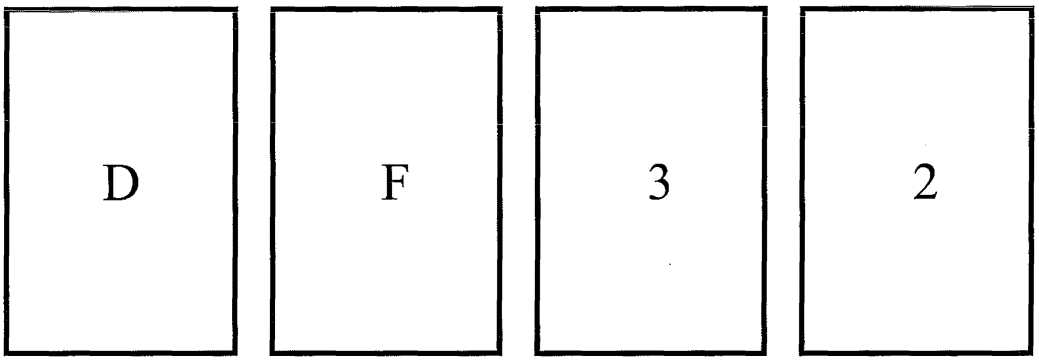
Note: Because of the rigid prescriptions of the University of Canterbury Library's thesis submission criteria, some of the following tasks have had to be reformatted to fit in to the required margins. Hence, they may appear 'squashed'. This was not the case for the tasks used in the experimental testing – all scenarios fitted comfortably on to one A4 sheet of paper without reformatting or compression.

Thematic-abstract Task:

You have just landed a clerical job at a local high school. Part of your job is to make sure that student documents have been processed correctly on the school database. Before you enter the data into the computer, you must check that the documents conform to the following alphanumeric rule:

"If a person has a 'D' rating, then his/her document must be marked code '3' ".

You suspect that the secretary you replaced did not categorise the students' documents correctly (apparently, the previous secretary was often unreliable). The cards below have information about the documents of four people who are enrolled at this high school. Each card represents one person. One side of a card tells a person's letter rating and the other side tells that person's number code. Indicate only those card(s) you definitely need to turn over to see if the documents of any of these people violates the rule above.



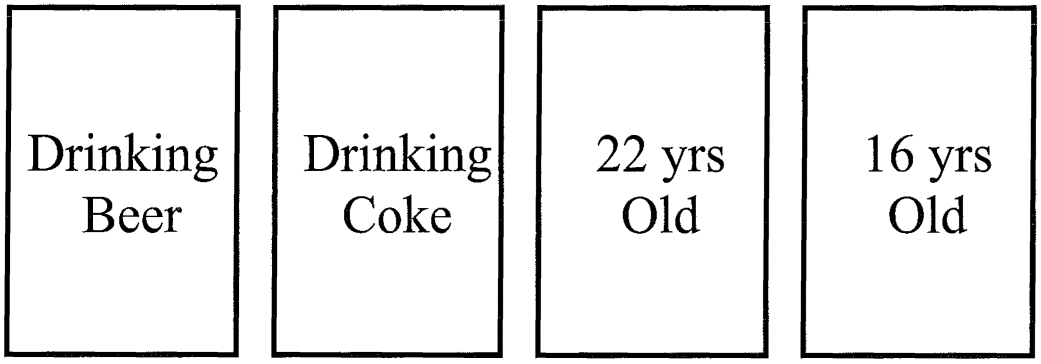
Drinking Age Rule (New Zealand adaptation):

You are a member of the local police force. It is your job to make sure people are conforming to certain rules. One of these rules relates to the consumption of alcohol: only people above a certain age are allowed to drink alcohol in a public place. This age has been arbitrarily set by the state as being 20 years of age. The rule states that only people over this age may frequent certain places such as pubs and bars. In other places, such as licensed restaurants, where people under this age are allowed, people may only drink alcohol if they are over 18 and consuming a meal as well.

It has recently been found that many establishments are flouting this law, so the police are imposing a crack-down on places that serve alcohol. As part of this crack-down you have been assigned to go around town checking bars and restaurants. The first establishment you enter is a restaurant. You see four young people sitting at a table. They are all drinking various beverages and eating meals. You go over to check that they are not breaking the law. The rule is:

"If a person is drinking beer, then the person must be over 18 years of age."

The cards below have information about each person. On one side is information about what age they are, and on the other side is information about what they are drinking. Indicate only those card(s) you definitely need to turn over in order to check whether any of the people are violating the rule.



Overnight Rule:

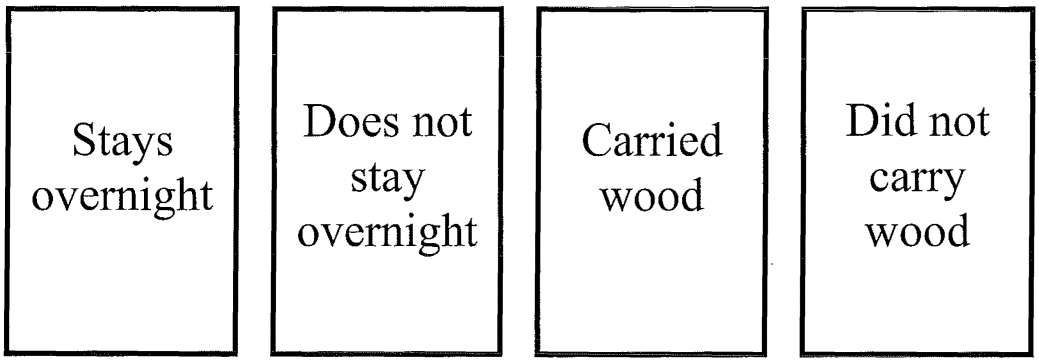
You are a park ranger for the Swiss Mountaineering Association. Among your many tasks as a ranger, you are charged with patrolling the Association's mountain cabins, making sure that they are kept in good order and that the hikers that use them follow the rules of the Association.

One of these rules relates to the transportation of firewood up to the cabins by the hikers. Although it is often warm enough in the cabins during the day when the sun is out, it is always cold at night, even during summer. Therefore, each cabin has a large fire place with which to heat it. The Association runs a lot of cabins but, unfortunately, it does not have enough resources to be able to keep a constant supply of firewood to all of them. Also, there is usually no wood around the cabins that could be used to build a fire, because most of the cabins are at high altitude where very few trees can grow. Therefore, the Association has a formal rule which all of it's members, or any other cabin users, must follow:

"If someone stays overnight in the cabin, then that person must bring along a bundle of wood from the valley below."

Use of the cabins is free to Association members, but non-members may use the cabins if they pay a small daily fee. The wood is also supplied by the Association at a depot in the valley at the bottom of the mountain; all the hikers have to do is pick it up and carry it to the cabin they will be staying overnight at. Unfortunately, some hikers still flout this rule. They would prefer to use wood that other hikers have carried up, without replacing the wood they use. Hikers who are only at a cabin for the day, and do not stay overnight, are also encouraged to bring wood, but it is not mandatory.

You have just reached one of the cabins situated about two hours hike above the tree-line. It is mid afternoon. There are four hikers resting in the cabin. They are all travelling individually, and are paid-up members of the Mountaineers Association. You decide to check to see if any of them are violating the rule above. Each card below represents information about a different hiker. One side of a card tells whether or not the hiker is staying overnight in the cabin, and the other side tells whether or not the hiker has carried a bundle of wood up the mountain from the valley depot. Indicate only those card(s) you definitely need to turn over to see if any hiker is violating the rule above.



Grover High School Scenario (New Zealand adaptation):

You supervise four parents who volunteered to help at the local Board of Education. Their job is to assign students from various suburbs to particular high schools according to which suburb they belong to. When you came into work today, you found the place a-buzz with rumor and innuendo. Each volunteer is the parent of a teenager who is about to enter high school, and each processed their own child's documents. Now rumors are flying that your volunteers cheated when assigning their own children to a school, in order to get them into a better school. Here is the situation:

Students are to be assigned either to Grover High School, located in the suburb of Grover, or to Hanover High School, located in the suburb of Hanover. Grover High is a great school with excellent facilities and a superior academic record. In contrast, Hanover High is a mediocre school with decrepit facilities and a low academic record.

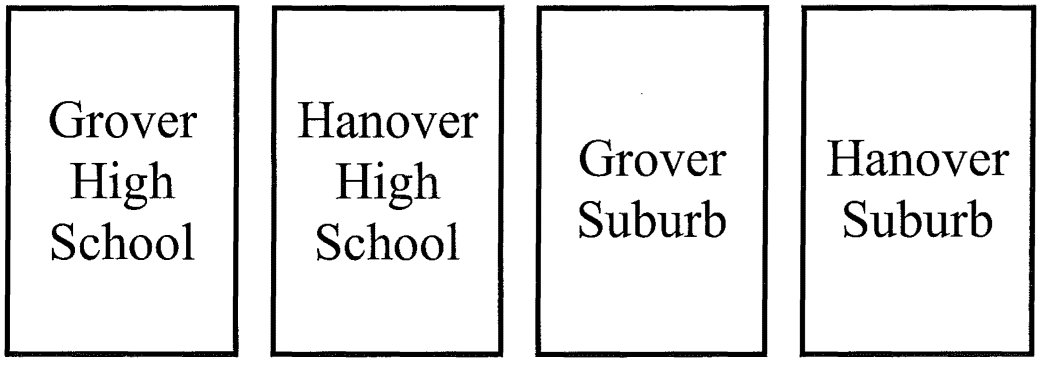
The reason the schools are so different is due to how willing the parents of each community are to financially support their schools through city rates and community funding. Although both communities are equally prosperous, the citizens of Grover have always cared about the quality of their schools, and have been willing to pay extra for it. In contrast, the parents in the neighboring suburbs of Hanover and Belmont never wanted to spend the money, and have opposed any rate-changes to improve Hanover High.

The Board of Education took these factors into account when it created rules to determine which school a student is to be assigned to. The most important of these rules is:

"If a student is to be assigned to Grover High School, then that student must live in the suburb of Grover".

Your volunteers were supposed to follow this rule when processing *all* students documents - including the documents of their own children! Most parents want to get their children the best education possible, however, some are not willing to pay for it. It is easy to imagine that your volunteers, being ambitious parents, might have been tempted to cheat on the rule. You task now is to find out whether any of them did:

The cards below have information about the documents of the four volunteers' children. Each card represents the child of one volunteer. One side of a card tells which school the volunteer assigned their child to, and the other side of the card tells what suburb the student lives in. Indicate only those card(s) you definitely need to turn over to see if the documents of any of these students violate the rule above.



Murder Investigation Scenario:

You are a police inspector investigating the homicide of a high-powered business executive. The facts of the case are thus: the victim collapsed in the car park of a restaurant that he had been dining in, and died before the ambulance arrived. A post-mortem indicated that the victim had been poisoned. The report also stated that the poison was ingested by the victim in white wine. You are treating the case as a murder inquiry.

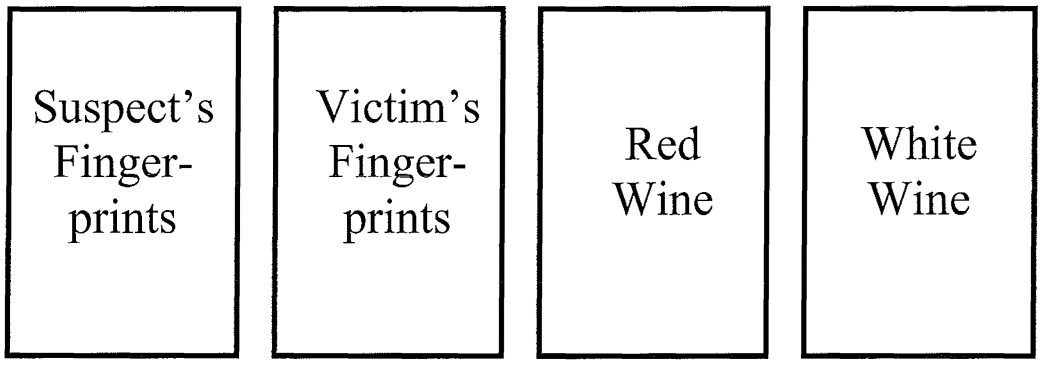
In order to investigate further, you decide to interview the prime suspect - a person that was dining with the victim at the restaurant just prior to the incident. This suspect was a work associate of the victim. The suspect had a possible motive for the murder because they are next in line for the victim's job (a rather powerful and prestigious one). On the other hand, the victim had a number of known business competitors who could benefit substantially from his demise, so there are other suspects.

The suspect claims that he and the victim had consumed a number of glasses of wine in the night with their meals. The waiter at their table that night confirms this; he had cleared up about a dozen wine glasses after the two had left. Also, all the wine had been ordered in glasses, not bottles. Luckily, none of the glasses had been washed before the police arrived, so the glasses could be analysed for fingerprints, poison and other evidence. Unfortunately, the glasses the two men had used were mixed up with the used glasses of other customers, so they could take a while to analyse properly.

The victim was seen going to the toilet about halfway through the dinner; you think that the suspect may have planted the poison in the victim's glass during this time. The suspect tells you that the victim drank both red and white wines that night. When the suspect is told that the poison was consumed in white wine, he suddenly claims "Ah, I couldn't have done it then ... I only ordered red wine that night and didn't touch the white wine". He is also adamant that he did not touch any of his partner's glasses. His body language suggests that he is lying. You decide to test his statement by turning it into the following form:

"If the suspect's finger-prints are on a glass, then that glass must have contained red wine."

The following "cards" below represent four types of wine glasses collected from the scene. On one side is shown information about whether the glass has the suspects finger-prints on it, the other side shows information about whether that glass contained red wine or white wine. Indicate only those card(s) you definitely need to turn over in order to check whether the suspect is lying or not?



Cookie Jar Caper Scenario:

You are the parent of a six year old boy who loves chocolate chip cookies. You keep the cookies in five glass jars which are kept in the pantry. You filled the jars only this morning and sealed them with large cork tops. But now some of the biscuits are missing and, as no one else has been in the house, you have a fair idea where they might have gone....

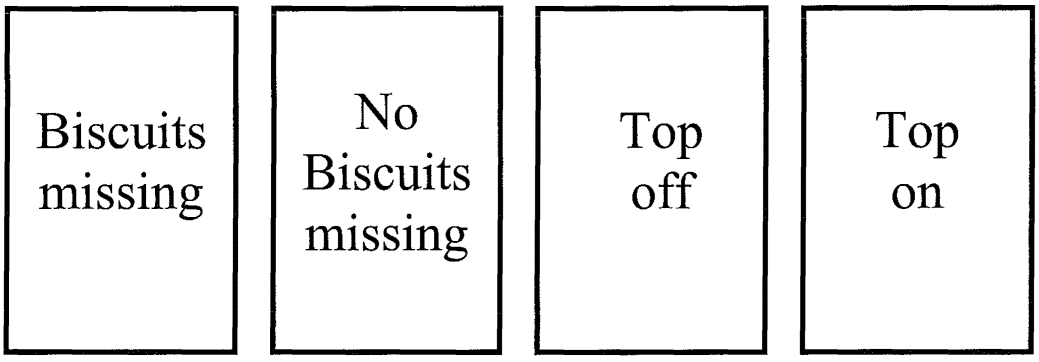
You first discovered the problem when you noticed the pantry door ajar. You went to close the door, then noticed that one of the cookie jars was lying on the floor with the top off. There were only a few biscuits left in it. Grumbling, you picked it up and put in back on the shelf with the other four. It was then that you noticed that some of the other jars were open, and that more of the cookies had gone.

Your son is playing in his bedroom with the family dog. You call him into the kitchen to explain himself. He looks at the jars and cries, "It wasn't me! It was the dog! I saw her!...". You rather suspect that he is lying, but you do have to admit that what he says *could* be the truth. The dog *does* like biscuits (it will eat just about anything, actually) and it has been known to steal food from the pantry when it thinks no one is looking. What you need is some way of checking to see whether your son is telling the truth.

You survey the crime scene again, looking for clues. You know that the dog could have pulled the tops off the jars with its teeth, but it could not possibly have put the tops back on again. Your son, of course, would have been capable of both actions. Hence, you can see one possible way of checking to see if your son is lying. You reason that if he is telling the *truth* (that is, the dog *did* do it) then the following statement must be true:

"If any biscuits are missing from a given jar, then its top must be off".

You decide to check whether the statement holds. The cards below represent the remaining four jars, each card a different jar. On one side of each card is information about whether the jar's top is on or off, and on the other side is information about whether any biscuits are missing from the jar. Indicate only those card(s) you definitely need to turn over in order to check whether your son is lying or not?



Missing Pharmaceuticals Scenario:

You are the manager of a manufacturing company which produces pharmaceuticals for sale at various pharmacies around the country and overseas. Your company has a superb record for reliability. Recently, though, a number of consumers and a consumer research group have indicated that there is a problem with one of your products. The problem is that this product, sold as tablets in small, sealed containers, is consistently found to be 'underweight' - that is, it there are not enough pills in some of the containers.

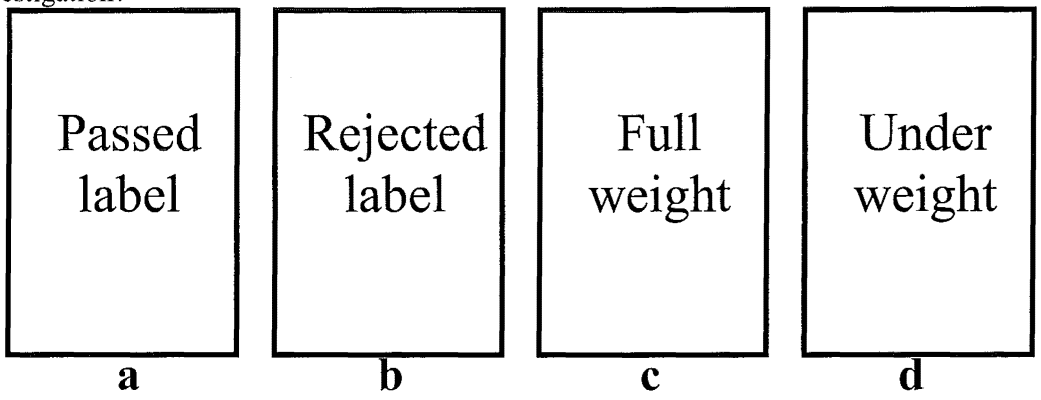
As with all manufactured products, you do expect some items to be defective. Normally, the number of underweight containers which escape detection is about 1 in 10,000, but the recent problem seems to have cropped up in 1 in every 10 containers sold. Your first thought was that the problem must have been caused by machinery or equipment fault, but rigorous testing has found that all the machines and equipment appear to be in perfect working order. Now you are beginning to suspect that the problem may have been due to theft by one of your workers.

The production process works like this. The tablets are made by a series of machines, then transported to a packaging machine that deposits the correct number of tablets into a container. The filled containers are then transported to a sorting machine. Each container is weighed to make sure that the contents are the correct weight. An incorrect weight could either mean that there are not enough pills in the container, or that some of the pills are too small. The machine then sorts each container according to whether or not it was the correct weight. A correct weight container is dropped down a chute labelled "full weight"; an incorrect weight container is dropped down a chute labelled "under weight". Any under weight containers are labelled "rejected" and put into a pile for later inspection. The full weight containers are then visually inspected by qualified staff for any defects or flaws. The containers that pass this inspection are sealed and labelled "passed". Any that fail the visual inspection are also labelled "rejected" and put in the same pile as the underweight containers.

It is at the stage of visual inspection that you suspect the problem is occurring. Because 1 tablet appears to be missing from every 10th container, you suspect that one of your ten inspectors may have been stealing the tablet before he or she seals the container with the "passed" label. The pills are quite expensive to produce and can be used to make narcotic drugs, hence they are highly sought after on the black market. Of course, to take the matter any further you are going to need proof of this. You decide an easy way of checking this is to take all the recently 'passed' and 'rejected' containers and put them through the sorting machine again. If the inspectors are doing their job correctly, then the following rule must be true:

"If a container is labelled 'Passed', then it must come out of the 'full weight' chute."

You reason that if this statement is found to be false, then you will have reasonable grounds to suspect theft by someone in the inspection staff, in which case you could investigate the matter further. The following cards represent four types of containers that have been fed back through the sorting machine. On one side of each card is information showing whether the container is labelled either 'passed' or 'rejected', and on the other side is information showing whether the container was sorted into the 'full weight' or 'under weight' chutes. Indicate only those card(s) you definitely need to turn over in order to check whether the inspection staff warrant further investigation?

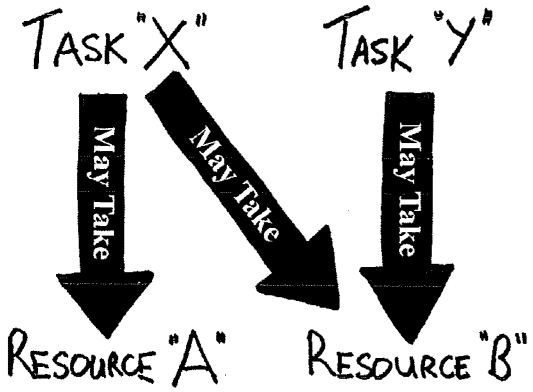


Appendix A2.2

Tasks for Experiment 2

In all examples "card a"= **P**,
"card b"=**~P**
"card c"= **Q**
"card d"=**~Q**.

Abstract Social Contract Scenario:



You are a member of certain community that relies on two particular resources (A and B) to ensure its members' survival. Both resources serve the same function, but one is considered more desirable then the other. Resource A is rare, potent and considered a coveted luxury. Resource B is common, weak and it is not highly prized.

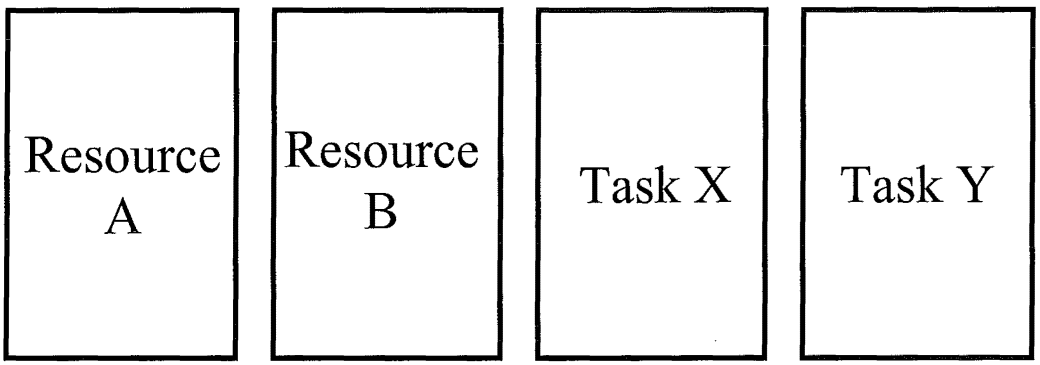
There are also a couple of tasks that each member of the community must perform in order to ensure the success of the community as a whole. The community member has the option of performing either task X or Y. Task X is a tough, risky task, but one which provides immense benefit for the community. Task Y is an easy, safe task, and is of only a small benefit to the community.

The community has come up with a law to limit the usage of Resource A and to encourage the participation in Task X. The law is:

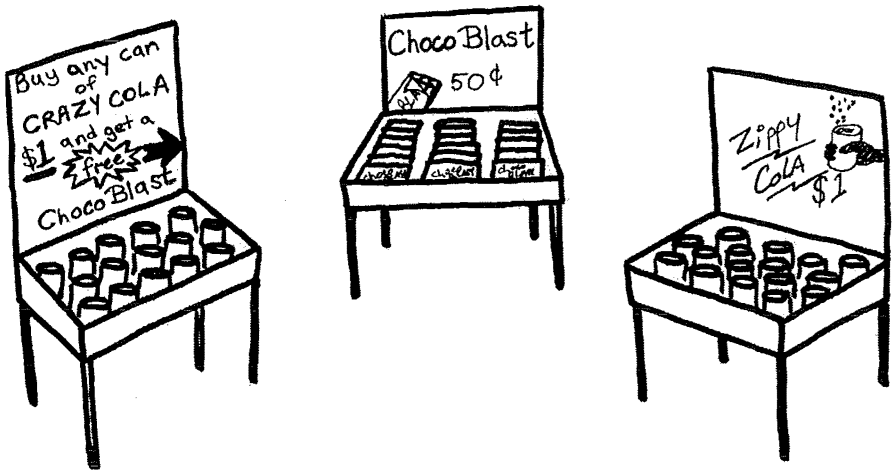
If a person uses resource 'A', then they must have performed task 'X'.

Your main job in the community is to enforce this law.

The cards below have information about four community members who have completed one of the tasks and used one of the resources. Each card represents one community member. One side of each card shows which task they performed and the other side shows which resource they used. Indicate only those card(s) you definitely need to turn over to see if any of these members have broken the law above.



Product Promotion Scenario:



You are the shopkeeper of a small, independent grocery store. Your store does good business and has a loyal local customer base. Occasionally, the manufacturers of the products you sell run special promotions to increase customer interest in their products. These promotions often include discounts, giveaways, competitions, or bundled sample products.

Currently, a particular soft-drink company is running a promotion on its "Crazy Cola" drink. The company is giving away a free "Choco-Blast" chocolate bar with every purchase of a can of "Crazy Cola" for $1:00. The free "Choco-Blast" offer is not available with any other soft drink brand, but customers may purchase the bars separately for 50 cents each.

The promotion works like this. Your store buys the soft drink and chocolate bars at the normal wholesale price. You then sell the "Crazy Cola" at the $1 retail price and let the customer have a "Choco-Blast" free. You keep a running total of the number of "Crazy Colas" you sell over the promotional period, and the soft drink company reimburses you the retail price for that many "Choco-Blasts" at the end. Hence, if someone buys a "Crazy Cola" but does not take a "Choco-Blast" you are still reimbursed for the chocolate bar. Likewise, if someone takes a "Choco-Blast" but does not buy a "Crazy Cola" then you do not get reimbursed (you will be out of pocket!)

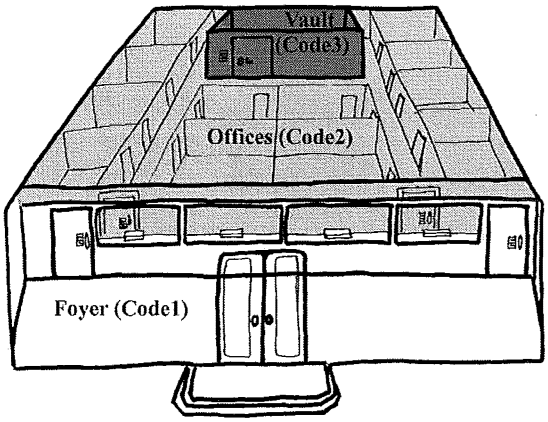
You suspect that some customers might try to take a "Choco-Blast" when they buy a can of "Zippy Cola", a more popular brand. To avoid this, you decide to pay special attention to anyone buying a can of cola. The rule of the promotion is:

"If a person takes a free "Choco-Blast", then they must buy a can of "Crazy Cola".

The cards below have information about four customers that are buying cans of cola. Each card represents one customer. One side of a card shows what brand of cola they are buying and the other side shows whether or not they have taken a free "Choco-Blast". Indicate only those card(s) you definitely need to turn over to see if any of these customers have broken the rule of the promotion.

Takes a free Choco-Blast	Doesn't take a free Choco-Blast	Buys Crazy Cola	Buys Zippy Cola
a	**b**	**c**	**d**

Security Guard Scenario:



You are a new security guard at a high-security government building. Many important documents are housed in the building such as records of the country's finances, personal and legal documents, and other confidential information. Many financial corporations, political groups, and unscrupulous individuals could benefit from this information, at the expense of the country as a whole. These factions have been known to bribe people who work in the building in order to get them to pass on these valuable documents.

Because of their sensitive nature, these documents are locked away within a large vault in the building. Only people with the correct security clearances are permitted into this vault. There are three levels of security. A 'Code 1' security pass only allows a person into the main foyer area that surrounds the vaults and offices, but not the offices or vaults themselves. A 'Code 2' security pass allows the person into the foyer and into the offices. A 'Code 3' security pass allows the person into the vault and into all the offices and the foyer.

It is your first day on the job. You are on night patrol. There is no one in the building except you, another security guard, and a few cleaners. All cleaners have been issued with at least a 'Code 2' security pass, so as they can clean up the main foyer and offices. But only a chosen few have been given a 'Code 3' pass. These chosen few are the longest serving and most trustworthy cleaners, who have been entrusted with the job of cleaning the vaults as well.

The other security guard is about to go on coffee break. He tells you that while he is away he wants you to check to see if any of the cleaners are breaking the security pass rule. He says the rule is:

'If a cleaner is in the Vault, then they must have a Code 3 security pass.'

The cards below contain information about four cleaners that are working in the building. Each card represents one cleaner. On one side of each card shows which area the cleaner is working in, and the other side of the card shows what sort of security pass the cleaner has. Indicate only those card(s) that you definitely need to turn over to see if any of the cleaners are breaking the rule.

Vault Area	Foyer Area	Code 3	Code 2
a	**b**	**c**	**d**

Engineering Diagram Scenario:

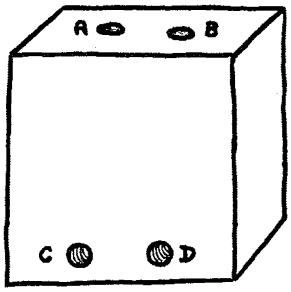


figure 1

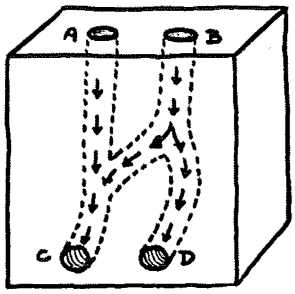


figure 2

You are an engineering student. For one of your labs you are given a box with two small holes in the top face and two small holes at the bottom of the front face (see figure 1). You are told that if a marble is dropped into one of the top holes (A or B) it will fall through a series of tubes before exiting out through one of the bottom holes (C or D). You can not look inside the box to see how the tubes are arranged.

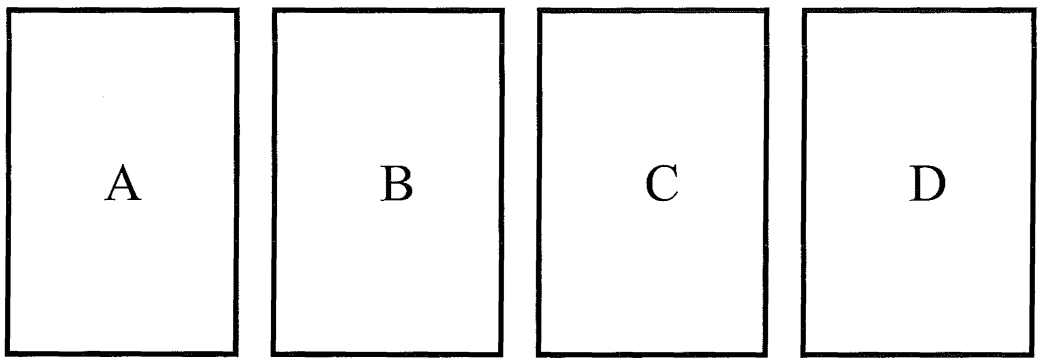
Next your lab supervisor gives you a diagram that shows one possible outline of the interior tubes (see figure 2). In this model, a marble dropped in hole A will always exit out hole C, but a marble dropped in hole B could either exit out hole C or hole D.

Your job is to test whether or not the diagram is correct by observing what happens when marbles are dropped down each hole. The diagram is also accompanied by a rule that describes what we would expect to find if the diagram is correct.

"If a marble goes in hole A, then it will always come out hole C."

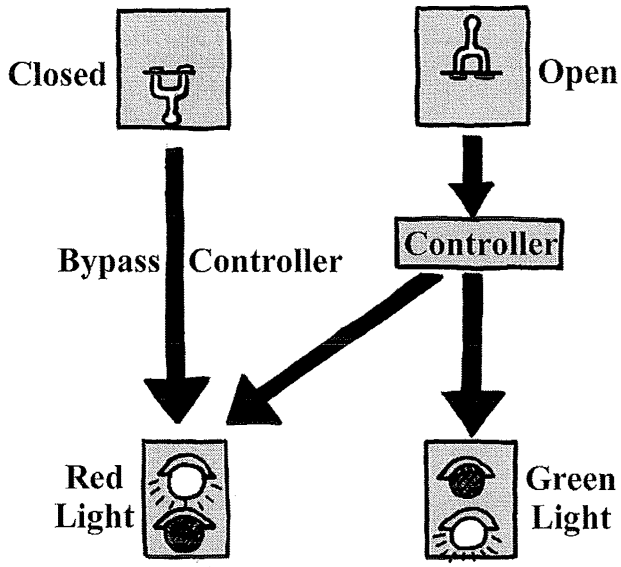
If this rule is wrong then the diagram must also be wrong.

The cards below contain information about four marbles that were dropped through the box. Each card represents one marble. On one side of each card shows which hole the marble was dropped into (A or B), and the other side of the card shows which hole the marble came out of (C or D). Indicate only those card(s) that you definitely need to turn over to see if the rule is wrong.



a **b** **c** **d**

Automated Rail Controller Scenario:



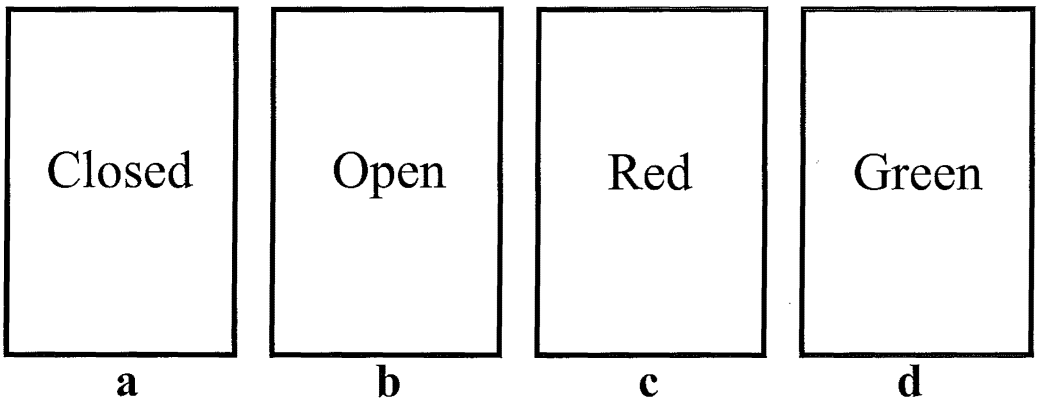
Recently Eastern Rail has invested in a new automated junction-controller for its railway lines. This computer-controlled system monitors the lines, and checks to make sure that all trains are running on the correct tracks. If the controller finds two trains heading towards each other on the same track, it changes the junctions to divert them onto separate lines. This is thought to be more efficient than manually monitoring the lines, and it eliminates human error. In the past such human error has caused a number of serious accidents and near misses.

Like all well designed safety systems, this automated junction-controller also has a number of "fail-safe" features. Firstly, it also controls a set of rail-traffic-lights on the lines. If one train is correctly heading down the tracks, the automated controller sets the lights to green to indicate that it is safe for the train-driver to proceed. If another train is found to be on the same lines then the automated controller changes the lights to red to indicate danger and the train-driver should stop. Secondly, there is a manual over-ride switch which sets all the traffic lights to red. This is used as a precautionary measure when it is suspected that the junction-controller is faulty.

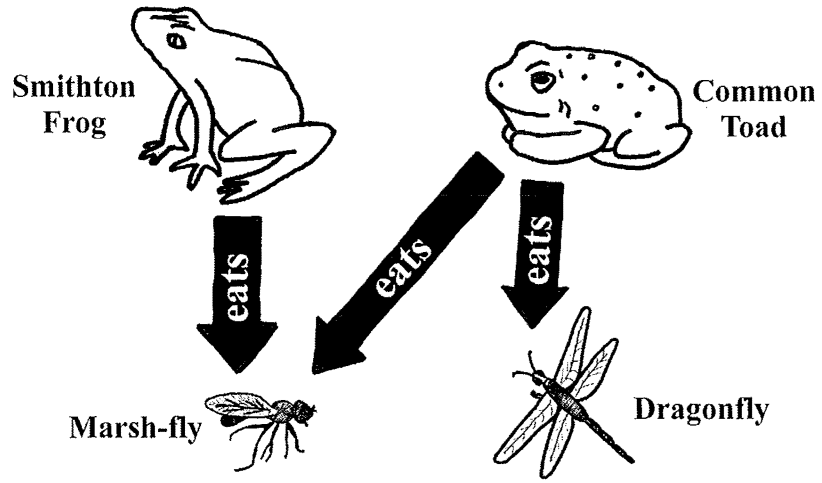
You are a safety engineer with eastern Rail. You maintain the system and check the safety features. Your job at the moment is to ensure that the manual over-ride switch is working properly. When the switch is OPEN the computer decides whether the traffic lights should be green or red, when the switch is CLOSED the computer is overridden and the lights turn red. Hence you are to check that the following rule is true:

"If the switch is CLOSED, the lights are RED".

The cards below contain information about four different tests on the junction-controller. Each card represents one test. On one side of each card shows whether the manual override switch was open or closed, and the other side of the card shows whether the traffic lights were red or green. Indicate only those card(s) that you definitely need to turn over to see if the manual override switch is working properly.



Smithton Swamp Scenario:



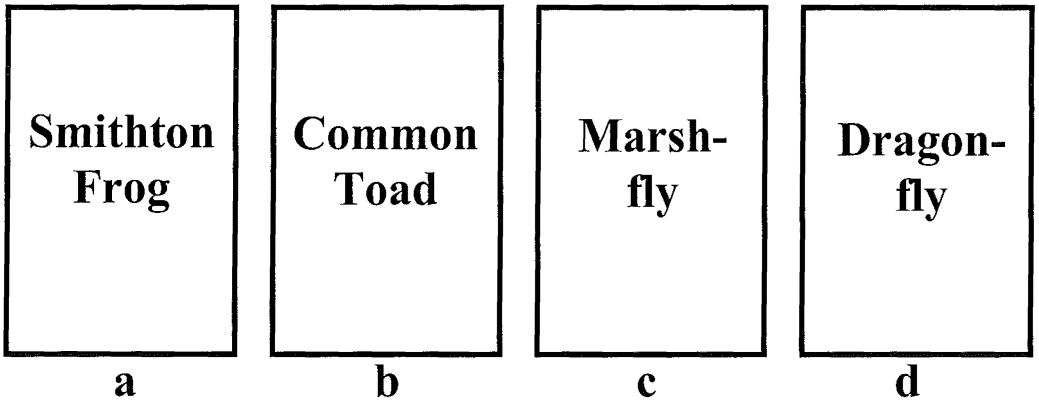
Smithton Swamp is a small wetland area on the outskirts of the city. For years it has been in danger of being drained as part of a housing development project. Conservationists have spent years trying to prevent this development. Until now, the local council has sided with the conservationists, but growth in the city has risen to such an extent that they feel that they have no real option but to go ahead with the development. Unfortunately, the swamp is home to a number of rare wetland insects and amphibians. Rather than simply eradicating the species, the local council is determined to relocate them to a wetland further out in the country. Although costly, the relocation posses no real problem for-all-but-one on the resident species.

This problem specie is a small frog that is known to only inhabit Smithton Swamp. While the other amphibian specie (a common toad) will eat all sorts of insects, the Smithton frog is thought to only eat marsh-flies. This fussiness is a problem because marsh-flies can not live in the new swamp because it is too acidic for them. Hence, although the Smithton frogs would be happy enough in the new swamp, they might starve to death. Dragonflies are quite populous at the new swamp, but it is not known whether the frogs will eat them. Marsh-flies do live in other swamps around the country, but moving the frogs to those parts would be far too costly.

You are an expert in species-relocation. You have been hired by the local council to determine whether the frogs will survive if relocated. You decide to study the feeding habits of the swamp's inhabitants and check the following theory:

If a Smithton frog is seen feeding, then it will only be eating a Marsh-fly.

The cards below contain information about four amphibians that you see feeding at Smithton Swamp. One side of each card shows what species the amphibian was, and the other side shows what the amphibian was eating. Indicate only those card(s) that you definitely need to turn over to see if the theory above is false.



Appendix A2.3

Information Sheets

Information Sheet for Exp 1a:

Information:

You are invited to participate as a subject in a research project aimed at investigating how well people reason under certain conditions.

Your involvement in this project will involve reading through a number of different stories or "scenarios" and, at the end of each one, answering a problem based on that scenario. I ask that you read each scenario carefully before answering (take your time, there's no rush!), and that you go through the problems in the order presented (without referring back to previous problems). I also ask that you do not confer with anybody else – we are only interested in *your* responses. These tasks should take around ten minutes to complete.

There are no risks to you in completing these tasks, and no deception is involved.

The results of the project may be published, but you may be assured of the complete confidentiality of the data gathered in this investigation: the identity of the participants will not be made public without their consent. To ensure anonymity and confidentiality the raw data will only be made accessible to the researcher and his supervisor.

This project is being carried out by Jonathan Grady, who can be contacted at 366-7001 Ext. 8084, under the supervision of Mr. Paul Russell (364-2170). Either person would be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and passed by the University of Canterbury Human Ethics Committee.

The Task:

You will have three short "stories" or "scenarios" to read through (one on each page). At the end of each scenario you will be asked to answer a problem which relates to the scenario. The problem will involve testing the truthfulness of a given **sentence**.

The **sentence** in question will be given in **bold** letters. To answer the problem there will be four "cards" given at the bottom of the page. These "cards" are depicted by rectangles which contain certain information. On each "side" of a card is information which could be used to test the truthfulness of the **sentence** in question. You will only get to see one side of each card.

To answer the problem you will be asked to indicate which card(s) (and which one(s) only) *must* be turned over in order to test the given **sentence**. To do this, circle the letter (a, b, c, or d) situated underneath the appropriate card.

Feel free to return to *this* page at any time to make sure you understand the instructions, **but** make sure that once you start a new problem, you do not refer to problems that you have already finished. Likewise, do not attempt a new problem until you have finished the one you are on.

Information Sheet for Exp 1b and Exp 2:

Information:

You are invited to participate as a subject in a research project aimed at investigating how well people reason under certain conditions.

Your involvement in this project will involve reading through a couple of different stories or "scenarios" and, at the end of each one, answering a problem based on that scenario. I ask that you read each scenario carefully before answering (take your time, there's no rush!), and that you go through the problems in the order presented (without referring back to the previous problem). I also ask that you do not confer with anybody else - we are only interested in *your* responses. These tasks should take around five-ten minutes to complete.

There are no risks to you in completing these tasks, and no deception is involved.

The results of the project may be published, but you may be assured of the complete confidentiality of the data gathered in this investigation: the identity of the participants will not be made public. To ensure anonymity and confidentiality the raw data will only be made accessible to the researcher and his supervisor.

This project is being carried out by Jonathan Grady, who can be contacted at 364-2987 Ext. 7986, under the supervision of Mr. Paul Russell (364-2170). Either person would be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and passed by the University of Canterbury Human Ethics Committee.

The Task:

You will have two short "stories" or "scenarios" to read through (one on each page). At the end of each scenario you will be asked to answer a problem which relates to the scenario. The problem will involve testing the truthfulness of a given **sentence**.

The **sentence** in question will be given in **bold** letters. To answer the problem there will be four "cards" given at the bottom of the page. These "cards" are depicted by rectangles that contain certain information. On each "side" of a card is information that could be used to test the truthfulness of the sentence in question. You will only get to see one side of each card.

To answer the problem you will be asked to indicate which card or cards (and which one(s) only) *must* be turned over in order to test the given **sentence**. To do this, circle the letter (a, b, c, or d) situated underneath the appropriate card or cards.

Feel free to return to *this* page at any time to make sure you understand the instructions, **but** make sure that once you start the second problem, you do not refer to the problem that you have already finished. Likewise, do not attempt the second problem until you have finished the first one.